

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGCTGGTAAGGATTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTGTGAAGTTGTAACAGAAAACCTGTTAGAAATGTGGTGGT
TTCAGCAAGGCCTCAGTTCCCTCAGCCCTGTAATTGGACATCTGCTGCTTCATATT
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGTTACCTTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTGGGCAATGCTAAATATTGCGGCAGTTTATGCATTG
CTACCATTATGTCGTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAA
TTAAACAAGGCTGGCCTGTACTTGAATACTGAGTTGTTAGGACTTCTATTGTGGCAAACCT
CCAGAAAACAACCCCTTTGCTGCACATGTAAGTGGAGCTGTGCTTACCTTGGTATGGCTCAT
TATATATGTTGTTGACGACCATCCTTCTACCAATGCAGCCCCAAATCCATTGCAACAAAGTC
TTCTGGATCAGACTGTTGTTATCTGGTGTGGAGTAAGTGCACCTAGCATGCTGACTTGCTC
ATCAGTTTGACAGTGGCAATTGGACTGATTAGAACAGAAACTCCATTGAAACCCGAGG
ACAAAGGTTATGTGCTCACATGATCACTACTGCAGCAGAATGGCTATGTCTTTCTTCTT
GGTTTCTGACTTACATCGTATTTCAGAAAATTCTTACGGGTGGAAGCCAATTACA
TGGATTAACCCCTATGACACTGCACCTGCCCTATTAACAATGAACGAACACGGCTACTTCCA
GAGATATTTGATGAAAGGATAAAATATTCTGTAATGATTGATTCTCAGGGATTGGGAAAGG
TTCACAGAAGTTGCTTATTCTCTGTAAATTCAACCACTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAGGCCATTGATAGATTATTCTAAAGGATATCAT
CAAGAAGACTATTAAAAACACCTATGCCTATACCTTATCTCAGAAAATAAGTCAAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIFSYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMOPKIHGKQVFIRLLLVIWCGVSALSMLTCSSVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFFLTYIRDfqKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCCGTGGGCGGACGCCGTGGGGAGAGCCGCAGTCGGGCTGCAGCACCTGGAGAACCCAGACC
GTGTGAGGGGGCCTGTGGCCCCAGCGTGTGGCCTCGGGAGTGGAAAGTGGAGGCAGGAGCCTTC
CTTACACTTCGCCATGAGTTCCTCATCGACTCCAGCATCATGATTACCTCCAGATACTATTTTG
GATTGGGTGGCTTCTTCATGCCAATTGTTAAAGACTATGAGATACTGTCAGTATGTTGTACAG
GTGATCTCTCCGTGACTTGCACATTCTGACCATGTTGAGCTCATCATCTTGAATCTGCTGATCCTGG
AGTATTGAATAGCAGCTCCGTTATTTCACTGGAAAATGAACCTGTGTGAATTCTGCTGATCCTGG
TTTCATGGCCTTTTACATTGGCTATTTATTGTGAGCAATATCGACTACTGCATAAACAAACGA
CTGCTTTTCTGCTCTTATGGCTGACCTTATGTATTCTCTGGAAACTAGGAGATCCCTTCC
CATTCTCAGCCAAAACATGGATCTTATCCATAGAACAGCTCATCAGCCGGTTGGTGTATTGGAG
TGACTCTCATGGCTCTTCTGGATTGGTGTCAACTGCCATACACTTACATGTCTTACTTC
CTCAGGAATGTGACTGACACGGATTCTAGCCCTGGAACGGCAGTGCTGCAAACCATGGATATGAT
CATAAAGCAAAAGAAAAGGATGGCAATGGCACGGAGAACAAATGTTCCAGAAGGGGAAGTGCATAACA
AACCATCAGGTTCTGGGAATGATAAAAAGTGTGTTACACTTCAGCATCAGGAAGTGAAAATCTTACT
CTTATTCAACAGGAAGTGGATGCTTGGAAAGAATTAGCAGGAGCTTCTGGAAACAGCTGATCT
ATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTCAAGGGGAATATTAAATTCTGGTT
ACTTTTCTCTATTACTGTGTTGGAAAATTTCATGGTACCATCAATATTGTTTGATCGAGTT
GGGAAAACGGATCCTGTACAAGAGGCATTGAGATCACTGTGAATTATCTGGAAATCCAATTGATGT
GAAGTTGGTCCCACACATTCCCTCATTCTGGAAATAATCATCGTCACATCCATCAGAGGAT
TGCTGATCCTTACCAAGTTCTTATGCCATCTTAGCAGTAAGTCCTCCAATGTCATTGCTCG
CTATTAGCACAGATAATGGCATGTACTTTGCTCCTCTGTGCTGCTGATCCGAATGAGTATGCC
AGAATACCGCACCATAACTGAAGTCCTGGAGAACACTGCAGTTCAACTCTATCACCGTTGGTT
ATGTGATCTTCTGGCAGCGCTCTCTAGCATACTCTTCTATTGGCTCACAAACAGGCACCA
GAGAAGCAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTCAAAATTAA
GATATAAGGGGGAAAATGGAACCAGGGCTGACATTATAACAAACAAAATGCTATGGTAGC
ATTTTCACCTTCATAGCATACTCCTCCCGTCAGGTGATACTATGACCATGAGTAGCATGCCAG
AACATGAGGGAGAGACTAATCAAGACAATACTCAGCAGAGAGCATCCGTGTGGATATGAGGCTGG
TGTAGAGGGAGAGGGAGCCAAGAAAATGAAACTAAAGGTAAAAACTGGAACCTGGGGCAAGACATGT
CTATGGTAGCTGAGCAAACACGTAGGATTCCGTTAAGGTCACATGGAAAAGGTTATAGCTTIG
CCTTGAGATTGACTCATTAAATCAGAGACTGTAACAAAAAAAAAAAAAGGGCGGCCGCG
ACTCTAGAGTCGACCTGCAGAAGCTGGCCGCCATGGCCAACCTGTTATTGCAGCTTATAATG

FIGURE 4

MSFLIDSSIMITSQILFFFGFWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV
LNSSSRYFHWKMNLCVILLILVFMPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFFWKLGD
FPILSPKHGILSIEQLISRGVIGVTLMALLSGFGAVNCPTYMSYFLRNVTDTDILALERRLQ
TMDMIISKKRMAMARRTMFQKGKYNFLGYFFSIYCVWKIFMATINIVFDRVGKDPVTRGIEI
LFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKDPVTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLAQIMGMY
FVSSVLLIRMSMPLYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGATGTCTCGTTATGAAGTGGAGCACTGAGTGTGAGCCTAACATAGTTCC
AGAACTCTCATCCGGACTAGTTATTGAGCATCTGCCTCTATACCACTGGCCATCTGAGGT
GTTTCCCTGGCTCTGAAGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTGCTTCTCACT
TCCATCTGGACCACGGCTCTGGTCAAGGCTTTGCGTGAGAAGAGCTTCCATCCAGGT
GTCATGAGAATTATGGGATCACCTTGAGCAAAAAGCGAACAGCAGCTGAATTTCACAG
AAGCTAAGGAGGCCGTAGGCTGCTGGGACTAAGTTGCCGGCAAGGACCAAGTTGAAACAGCC
TTGAAAGCTAGCTTGAAACTTGAGCTATGGCTGGGTTGGAGATGGATTGCGTGTATCTCTAG
GATTAGCCCACACCCAAGTGGGGAAAATGGGGTGGGTGCTGATTGAAAGGTTCCAGTGA
GCCGACAGTTGAGCCTATTGTTACAACCTATCTGATACTTGGACTAACTCGTGCATTCCAGAA
ATTATCACCAACCAAAGATCCCATATTCAACACTCAAACACTGCAACACAAACAGAATTATGT
CAGTGACAGTACCTACTCGGTGGCATCCCTTAECTCTACAATACCTGCCCTACTACTCTC
CTGCTCCAGCTTCCACTTCTATTCACGGAGAAAAAAATTGATTGTTGTCACAGAAGTTTATG
GAAACTAGCACCATGTCTACAGAAACTGAACCATTGTTGAAAATAAGCAGCATTCAAGAATGA
AGCTGCTGGTTGGAGGTGTCACGGCTCTGCTAGTGCTCTCCTCTTGGTGTG
CAGCTGGCTTGAGTTTGCTATGTCAAAGGTATGTGAAGGCCCTCCCTTACAAACAAGAAT
CAGCAGAAGGAAATGATCGAAACCAAAGTAGTAAAGGAGGAGAACGCAATGATAGCAACCCCAA
TGAGGAATCAAAGAAAATGATAAAAACCCAGAACAGAGTCAAGAGTCAAGCAAACCTACCGTGC
GATGCCTGGAAGCTGAAGTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTCTT
CATGCTCCTTACCCCTGCCAGCTGGGAAATCAAAGGCCAAAGAACCAAAGAACGAAAGTCCA
CCCTGTTCTAACTGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAAC
GCCCTCTCTTATTGTAACCTGTCTGGATCCTATCCTCCTACCTCAAAGCTTCCACGGCCT
TTCTAGCCTGGCTATGCTTAATAATATCCACTGGGAGAAAGGAGTTTGCAAGTGCAAGGAC
CTAAAACATCTCATCAGTATCCAGTGGTAAAAGGCCCTCTGGCTGCTGAGGCTAGGTGGTTG
AAAGCCAAGGAGTCACTGAGACCAAGGCTTCTCTACTGATTCCGAGCTCAGACCCCTTCTCA
GCTCTGAAAGAGAAACACGTATCCCACCTGACATGTCCTCTGAGCCGGTAAGAGCAAAGAAC
GGCAGAAAAGTTAGCCCCCTGAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTCTGTA
GCTAAAATAAGAAATAGAACAGGCTGAGGATACGACAGTACACTGTCAGCAGGGACTGTA
ACAGACAGGGTCAAAGTGTCTCTGAACACATTGAGTTGGAATACTGTTAGAACACACACA
CTTACTTTCTGGTCTCTACCACTGCTGATATTCTAGGAATATACTTTACAAGTAACA
AAAATAAAAATCTTATAAAATTCTATTCTGAGTTACAGAAATGATTACTAAGGAAGATT
ACTCAGTAATTGTTAAAAGTAATAAAATTCAACAAACATTGCTGAATAGCTACTATATGTC
AAAGTGTGCAAGGTATTACACTCTGTAATTGAATATTCTCCTCAAAAATTGACATAGTAG
AACGCTATCTGGAGCTATTGTTGATATTCTAGCTATCTACTTCCAAACTAAT
TTTTATTCTGAGACTAATCTTATTCTAATATGGCAACCATTATAACCTTAAATT
TATTATTAAACATAACCAAGAAGTACATTGTTACCTCTATATACCAAAAGCACATTAAAAGTGC
ATTAACAAATGTACTAGCCTCTTCCAAACAAGAAGGACTGAGAGATGCAGAAATATT
TGTGACAAAAAATTAAAGCATTAGAAAACCTT

50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

FIGURE 6

MARCFSLVLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTAKEACRLLG
LSLAGKDQVETALKASFETCSYGVGDGFVVISRISPNCGKNGVGVLIWKVPVSRQFAAYCYN
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPPAPASTSIPR
RKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLALLFFGAAAGLGFCYVK
RYVKAFFTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCGGCCTCCGCACCCGGCCCCACCGGCCGCTCCGCATCTGCACCCGCAGCCCCGC
GGCCTCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCGAGCGCAACTCGGTCCAGTCGGGGCG
CGGCTGGGGCGCAGAGCGGAGATGCAGCGGCTTGGGGCACCCGTGCTGTGCCTGCTGGCG
CGCGGTCCCCACGGCCCCCGCGCCGCTCCGACGGCACCTCGCTCCAGTCAGCAAGCCGGCCG
GCTCTCAGCTACCCGCAGGAGGCCACCCCAATGAAGATGTTCCCGAGGGTTGAGGAACGTGAT
GGAGGACACGCAGCACAAATTGCGCAGCGGGTGGAAAGAGATGGAGGCAGAAGAAGCTGCTGCTA
AACATCATCAGAAGTGAACCTGGAAACTTACCTCCAGCTACAAATGAGACCAACACAGAC
ACGAAGGTTGAAATAATACCATCCATGTGCACCGAGAAATTACAAGATAACCAACAACCAGAC
TGGACAAATGGTCTTCAGAGACAGTTACACATCTGTTGGAGAGCAAGAAGGAGAAGGAGCC
ACGAGTGCATCATCGACGAGGACTGTGGGCCAGCATGTAUTGCCAGTTGCCAGCTTCCAGTAC
ACCTGCCAGCCATGCCGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAGTGCTGTGGAGACCA
GCTGTGTGCTGGGGTCACTGCACCAAAATGCCACCAAGGGGAGCAATGGGACCATCTGTGACA
ACCAGAGGGACTGCCAGCCGGGCTGTGCTGTGCCCTCAGAGAGGGCTGCTTCCCTGTGTGC
ACACCCCTGCCCGTGGAGGGAGCTTGCATGACCCCGCCAGCCGGCTCTGGACCTCATCAC
CTGGGAGCTAGAGCCTGATGGAGCCTGGACCGATGCCCTGTGCCAGTGGCCTCTGCCAGC
CCCACAGCCACAGCCTGGTGTATGTGCAAGCCGACCTTGTGGGGAGCCGTGACCAAGATGG
GAGATCCTGCTGCCAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCAATGGAGGAGGTGCG
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCCTGGGGAGCCTGCCAGT
CCGCCGCTGCACTGCTGGAGGGAAAGAGATTAGATCTGGACCAGGCTGTGGTAGATGTGCAA
TAGAAATAGCTAATTATTCAGGTGTGCTTAAAGGCTGGGTGACCAGGCTTCTTCTA
CATCTCTCCAGTAAGTTCCCCTCTGGCTTGACAGCATGAGGTGTGCAATTGTCAGCT
CCCCCAGGCTGTTCTCCAGGCTCACAGTCTGGCTGGAGAGTCAGGCAGGGTTAAACTGCA
GGAGCAGTTGCCACCCCTGTCCAGATTATGGCTGCTTGCCCTACAGTTGGCAGACAGCCG
TTTGTCTACATGGCTTGATAATTGTTGAGGGAGGAGATGAAACAATGTGGAGTCTCCCTC
TGATTGGTTTGGGAAATGTGGAGAAGAGTGCCTGCTTGCACACATCAACCTGGCAAAATG
CAACAAATGAATTTCACGCAGTTCTCATGGGATAGGTAAAGCTGTGCCTCAGCTGTG
AGATGAAATGTTCTGTTCACCTGCATTACATGTGTTATTCACTCAGCAGTGTGCTCAGCTCC
TACCTCTGTGCCAGGGCAGCATTCTACATCCAAGATCAATTCCCTCTCAGCACAGCCTGGG
AGGGGTCTTGTCTCCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTGCC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTCATCTGGTTGTGACTCTAAGCTCAGTGC
CTCCACTACCCCACACCAGCCTGGTGCACCAAAAGTGCCTCCAAAAGGAAGGAGAATGGGAT
TTTCTTGAGGCATGCACATCTGAATTAGGTCAAACATCTCACATCCCTCTAAAGTAA
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGGCAGCCGCTCTCTAATGAAGACAATGAT
ATTGACACTGTCCTCTTGGCAGTTGCATTAGTAACATTGAAAGGTATATGACTGAGCGTAGCA
TACAGGTTAACCTGCAGAAACAGTACTTAGTAATTGTAGGGCGAGGATTATAAATGAAATTG
AAAATCACTTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC
TGTGTGAAACATGGGTGTAATATGCACTGCGAACACTGAACCTACGCCACTCCACAAATGATG
TTTCAGGTGTGACTGGACTGTTGCCACCATGATTCACTCAGAGTTCTTAAAGTTAAAGTTGCA
CATGATTGTATAAGCATGCTTCTTGAGTTAAATTATGTATAAACATAAGTTGCAATTAGAA
ATCAAGCATAAATCACTCAACTGCAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL
RSAVEEMEAEEAAKASSEVNLPPSYHNETNTDTKVGNNTIHVREIHKITNNQTGQMVFSE
TVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRQRMILCTRDSECCGDQLCVWGHC
TKMATRGSGNTICDNQRDCQPGLCCAFQRGLFPVCTPLPVEGELCHDPASRLLLITWELEPDG
ALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGIELLPREVPEYEVGSFMEEVRQELEDLE
RSLTEEMALGEAAAAALLGGEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACCGCTGGCGGACCGTGGGGCTGTGAGAAAGTGCCAATAAACATCATGCAACCCAC
GGCCCACCTTGTGAACCTCTCGTGCCAGGGCTGATGTGCGTCTCCAGGGCTACTCATCAAAG
GCCTAATCCAACGTTCTGTCTCAATCTGCAAATCTATGGGTCTGGGCTCTCTACTGGCCTT
AACTGGTACTGGCCTGGCCATGCGTCTCGCTGGAGCCTTGCCCTCTACTGGCCTT
CCACAAGCCCCAGGACATCCCTACCTTCCCTTAATCTCGCCTCATCCGACACTCCGTACC
AACTGGTCATTGGCATTGGAGCCCTCATCCTGACCCCTGTGCAGATAGCCGGTCATCTG
GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCCTGTAGCCCGCTGCATCATGTGCTGTT
CAAGTGCTGCCCTGGTCTGGAAAAATTATCAAGTTCTAAACCGCAATGCATACATCATGA
TCGCCATCTACGGGAAGAATTCTGTGTCTCAGCCAAAATGCGTCATGCTACTCATGCGAAC
ATTGTCAGGGTGGTCGTCTGGACAAAGTCACAGACCTGCTGCTGTTCTGGAGCTGCTGGT
GGTCGGAGGCCTGGGGCTCTGTCTTCTTCTCCGGTCGCATCCGGGCTGGTAAAG
ACTTTAAGAGCCCCCACCTCAACTATTACTGGCTGCCATCATGACCTCCATCTGGGGCCTAT
GTCATGCCAGCGGTTCTCAGCGTTTCCGGATGTGTGGACACGCTTCCCTGCTTCT
GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGCCCTACTACATGTCCAAGAGCCTCTAA
AGATTCTGGCAAGAAGAACGAGGCGCCCCCGACAACAAGAAGAGGAAGAAGTGACAGCTCCGG
CCCTGATCCAGGACTGCACCCCACCCACCGTCCAGCCATCCAACCTCACCTCCCTACAGGT
CTCCATTTGTGGTAAAAAAAGGTTTAGGCCAGGCCTGGCTCACGCCGTGAATCCAACACT
TTGAGAGGCTGAGGCCGGCGGATCACCTGAGTCAGGAGTTGAGACAGCCTGCCAACATGGT
AAACCTCCGTCTATTAAAAACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCACTCCCA
GCTACTCGGGAGGCTGAGGCAGGAGAATCGCTGAAACCCGGAGGCAGAGGTTGCAGTGAGCCGA
GATCGGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAAA
AAGATTTATTAAAGATATTGTTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPTAHVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLFWTL
NWVLALGQCVLAGAFASFYWAFHKPQDIPFPLISAFIRTLRYHTGSIAFGALILTLVQIARVIL
EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAFMILLMRN
IVRVVVLDKVTDLFFFGLVVGVGVLSEFFFFSGRIPGLGKDFKSPHLNYWLPIMTSILGAY
VIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYYMSKSLLKILGKKNEAPPDNKKRK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

GCCCCGGCCCGGCCGGGGCCCCGAAGCCGGGAGCCACCGCATGGGGCTGCCTGGGAGCCTGC
TCCCTGCTCAGCTGCCGTCTGCCCTGCCGTCTGCCCTGCATCCITGCAGCTGCTGCCCGC
CAGCCGCAACTCCACCGTGAGCCCTCATCTCACGTTCTTCCTCTTCCGGGGTGTGGTGTCCA
TCATTATGCTGAGCCGGCGTGAGAGTCAGCTACAAAGCTGCCCTGGGTGTGAGGAGGGGCC
GGGATCCCCACCGTCCCTGCAGGCCACATCGACTGTGGCTCCCTGCTTGGCTACCGCCTGTCTACCG
CATGTGCTCGCCACGGCGGCCCTCTTCTTCCACCCGTCTCATGCTCGCTGAGCAGCA
GCCGGGACCCCCGGGCTGCCATCCAGAATGGGTTGGTCTTAAGTCCGTACCTGGTGGCCTC
ACCGTGGGTGCCTCTACATCCCTGACGGCTCCTCACCAACATCTGGTCTACTTCGGCGTGG
CTCCTCTCTCATCCCATCCAGCTGGTGTCTCATCGACTTGCCTACCTGGAACCCAGCGGT
GGCTGGGCAAGGCCGAGGAGTGCATTCCGTGCCTGGTACGCAGGCCCTTCTTCACTCTC
TTCTACTTGCTGTCATCGCGCGTGGCGTGTGTTACTACACTGAGGCCAGCGCTGCCA
CGAGGGCAAGGTCTTCATCAGCCTCAACCTCACCTCTGTGTCTGCCTGCTGCTGCTGC
CCAAGGTCCAGGACGCCAGCCAACCTGGCTGCTGCAGGCCCTCGGTACCATCCCCTACACCATG
TTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCCATTGCCAACCCAGCT
GGGCAACGAGACAGTTGTGGCAGGCCCGAGGGCTATGAGACCCAGTGGTGGGATGCCCGAGCATG
TGGGCCTCATCATCTCCTCTGTGCACCCCTTCTCATCAGTCTGCCTCAGACCACCGCAGGTG
AACAGCCTGATGCAGACCGAGGAGTGCCTACCTATGCTAGACGCCACACAGCAGCAGCAGCAGG
GGCAGCCTGTGAGGGCCGGCTTGACAACGAGCAGGCCCTGACAGCTACAGTACTCCTCTCC
ACTTCTGCCTGGTGTGGCTCAGTCACGTCATGATGACGCTACCAACTGGTACAAGCCGGTGAG
ACCCGGAAGATGATCAGCACGTGGACCGCCGTGTGGTGAAGATCTGTGCCAGCTGGCAGGGCTGCT
CCTCTACCTGTGGACCCCTGGTAGGCCACTCCTCTGCCAACCGCAGCTCAGCTGAAGGCAGCCTCA
CAGCCTGCCATCTGGTGCCTCTGCCACCTGGTGCCTCTGGCTCGGTACAGCCAACCTGCCCTC
CCCACACCAATCAGCCAGGCTGAGCCCCCACCCCTGCCAGCTCCAGGACCTGCCCTGAGCCGGC
CTTCTAGTCGTAGTCCCTCAGGGTCCGAGGAGCATCAGGCTCTGCAGAGCCCCATCCCCCGCCAC
ACCCACACGGTGGAGCTGCCCTTCCTCCCTCCCTGGTGCCTACTCAGCATCTGGATGAA
AGGGCTCCCTGTCTCAGGCTCCACGGGAGCAGGGCTGCTGGAGAGAGCAGGGAACTCCCACACAG
TGGGGCATCCGGCACTGAAGCCCTGGTGTCTGGTACAGTCCCCCAGGGACCCCTGCCCTTCTG
GACTTCGTGCCCTACTGAGTCTCTAAGACTTTCTAATAAAACAAGCCAGTGCCTGTAFFFFFF

FIGURE 12

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVESQL
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFTLLMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILQLVLLIDFAHSWNQRWLKGAE
ECDSRAWYAGLFFFITLLFYLLSIAAVALMFYYTEPSGCHEGKVFISLNLTFCVCVSIAAVLPKV
QDAQPNSGLLQASVITLYTMFTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI
VGLIIIFLLCTLFISLRSSDHRQVNLSMQTEECPPMLDATQQQQQQVAACEGRAFDNEQDGVTYSY
SFFHFCLVLASLHVMMTLTNWYKPGTRKMISTWTAVWKICASWAGLLLWTLVAPLLRNRD
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGCCGGCCAGGAACCACCGTTAAGGTGTCTCTCTTAA
GGTTGGAAAAAGACTCCTGTAACCCCTCCAGGATTGAACCACCTGCCAGAAGACATGGAGAACG
CTCTCACCGGGAGCCAGAGCTCCATGCTCTGCGCAATATCCATTCCATCAACCCACACAA
CTCATGGCAGGATTGAGTCCTATGAAGGAAGGGAAAAGAAGGCATACTGATGTCAGGAGGAC
TTTCTGTTGTTGTCACCTTGACCTCTTACCGTAAACATTACTGTGGATAATAGAGTTAAATG
TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTCATAT
TTTGATATATTCTCTGGCAGTTTCGATTTAAAGTGTAAACTTGCAATGCTGTGCAG
ACTGCGCCATTGGTGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTACTAGCAAAG
TGATCCTTCGAAGCTTTCTCAAGGGCTTGCTATGTGCTGCCATCTTCAATTCTCATC
CTTGCCTGGATTGAGACGTGGTCTGGATTCAAAGTGTACCTCAAGAACAGAAGAAGAAAA
CAGACTCCTGATAGTCAGGATGCTTCAGAGAGGGCAGCACTTACCTGGTGGTCTTGATG
GTCAGTTTATTCCCTCCTGAATCCGAAGCAGGATCTGAAGAACGCTGAAGAAAAACAGGACAGT
GAGAAACCACTTTAGAACTATTGAGTAACCTTTGTAAATGTGAAAAACCCCTCACAGAAAGTC
ATCGAGGCAAAAGAGGCAGGCAGTGGAGTCTCCCTGTCGACAGTAAAGTTGAAATGGTACGTC
CACTGCTGGCTTATTGAACAGCTAATAAGATTATTATTGTAAACCTCACAAACGTTGTAC
CATATCCATGCACATTAGTTGCCTGCCTGCGCTGGTAAGGTAATGTCATGATTCACTCTCT
TCAGTGAGACTGAGCCTGATGTGTTAACAAATAGGTGAAGAAAGTCTGTGCTGTATTCTAATC
AAAAGACTTAATATATTGAAGTAACACTTTTAGTAAGCAAGATAACCTTTATTCAATTAC
AGAATGGAATTTTTTGTTCATGTCAGATTATTGTATTCTTTAACACTCTACATT
TCCCTGTTTTAACTCATGCACATGTGCTCTTGTACAGTTAAAAGTGTAAATAAAATCG
ACATGTCAATGTGGCTAGTTTATTTCGATTATGTGTATGGCCTGAAGTGTGG
CTTGCAAAAGGGAGAAAGGAATTGCGAATACATGTAAAATGTCACCAGACATTGTATTATT
TTATCATGAAATCATGTTCTGATTGTTCTGAAATGTTCAAATACTCTTATTGAAATGC
ACAAAATGACTTAAACCATTCATGTTCCCTTGCCTCAGCCAATTCAATTAAAATGAA
CTAAATAAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLF
VTLLWIIELNVNGGIENTLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT
AVTSAFLLAKVILSKLFSQGAFGYVLPPIISFILAWEITWFLDFKVLQPQEAEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEEAEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACCGCAGTTGCTTCGGGACCCAGGACCCCCCTGGGCCCGACCCGCAGGAAAGACTGAGG
CCGCGGCCCTGCCCGCCCGCTCCCTGCGCGCCGCCCTCCGGGACAGAAGATGTGCTCCAG
GGTCCCCTCTGCTGCTGCCGCTGCTCCCTGCTACTGGCCCTGGGGCTGGGGTGCAGGGCTGCCCAT
CCGGCTGCCAGTGCAAGCAGCACAGACTTCTGCACTGCCGCCAGGGGACCACGGTGCC
CGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTGAGAACGGCATACCATGCTGACGC
AGGCAGCTTGGCCCTGCGGGCTGCACTGGACCTGTACAGAACAGATGCCAGCC
TGCCCAGCGGGCTTCCAGCACTGCCAACCTCAGCAACCTGGACCTGACGGCAACAGGCTG
CATGAAATACCAATGAGACCTCCGTGGCTGCGGCCCTCGAGGCCCTACCTGGCAAGAA
CCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGCTCGACCGCCTCTGGAGCTAAGCTG
AGGACAACGACTGCCGCACTGCCCTGGCTGCGCTGCCCGCTGCTGCTGGACCTCAGC
CAAACAGCCTCTGGCCCTGGAGGCCGATCTGGACACTGCCAACCTGGAGCGCTGCCG
GGCTGGCTGGGGCTGCACTGGAGGGCTCTCAGCGCTTGCGCAACCTCCACGACC
TGGATGTCGTCGACAACCAGCTGGAGCGAGTGCCACCTGTATCCGAGGCCCTCCGGGCTGACG
CGCCTGCCGCTGGCCGAACACCCGATTGCCAGCTGCCGCCAGGACCTGCCGCCCTGCC
TGCCCTGAGGAGCTGGATGTGAGCAACCTAAGCCTGCAAGGCCCTGCCCTGGCGACCTCTCGGCC
TCTTCCCCGCCCTGCGCTGCTGGCAGCTGCCGCAACCCCTCAACTGCGTGTGCCCTGAGC
TGGTTGGCCCTGGGTGCGCAGAGGCCACGTACACTGGCCAGGCCCTGAGGAGACGCGCTGCCA
CTTCCGCCAAGAACGCTGGCCGCTGCTCTGGAGCTTGACTACGCCGACTTGGCTGCCAG
CCACCACCAACAGCACAGTGCCAACACAGGCCCCGTGGTGCAGGCCACAGCTTGCT
TCTAGCTTGGCTCTACCTGGCTTAGCCCCACAGGCCGCAACTGAGGCCAGCCG
CACTGCCAACCGACTGTAGGGCTGTCCCCCAGCCCCAGGACTGCCAACCGTCAACCTGCCCTCA
ATGGGGCACATGCCACCTGGGACACGGCACCCACTGGCGTGTGGCCCGAAGGCTTCACG
GGCTGTACTGTGAGAGGCCAGATGGGGCACGGGACACGCCAGGCCCCCTACACCAGTCACGCCAG
GCCACCAACGGCTCCCTGACCTGGCATCGAGCCGGTGAAGCCCCACCTCCCGCCTGGGCTG
AGCGCTACCTCCAGGGGAGCTCGTGCAGCTCAGGAGCTCCGTCTCACCTATCGCAACCTATCG
GGCCCTGATAAGCGGTGGTGCAGCTGCCACTGCCCTGCTCGCTGAGTACACGGTACCCA
GCTCGGCCAACGCCACTTACTCGTCTGTGTCATGCCCTTGGGCCCGGGCTGGGAG
GCGAGGAGGCTGCCGGGAGGCCATACACCCCAAGCCGCTCACTCCAACACGCCAGTCACC
CAGGCCGCGAGGCCACCTGCCCTCTCATTGCGCCGCCCTGCCCGCGGTGCTCTGCCGC
GCTGGCTGCCGGTGGGGCAGCTACTGTGTGCGGGGGGGGGCATGGCAGCAGCGGCTCAGG
ACAAAGGGCAGGTGGGCCAGGGCTGGGACTGGAGGGAGTGAAGGTCCCCCTGGAG
CCAGGCCGAAGGACAAGAGGGCTGGGACTGGGCCCTGCCAGCGGTCTGAGTGTGAGGTGCC
ACTCATGGCTTCCAGGGCTGGGCTCCAGTCACCCCTCCACGCAAAGCCCTACATCTAAGCCA
GAGAGAGACAGGGCAGTGGGCCGGGCTCTCAGCCAGTGAGATGCCAGCCCCCTCTGCTGCC
ACACCACGTAAGTCTCAGTCCAAACCTGGGGATGTGTCAGACAGGGCTGTGTGACCACAGCT
GGGCCCTGTTCCCTCTGGACCTCGCTCCATCTGTGAGATGCTGTGGCCAGCTGACGCC
CTAACGTCCCCAGAACGGAGTGCTATGAGGACAGTGCTCCGCCCTGCCCTCCGCAACGTGCA
CCTGGCGACGGCGGCCCTGCCATGTGCTGTAACGCACTGCCCTGGGTGCTGGCTCTCCAC
TCCAGGCCAGGGCCACTGAAGGAAGCTCCCGAAAGAGCAGAGGGAGAGCGGGTAGGC
GGCTGTGACTCTAGTCTTGCTGGCCCAAGGAAGCGAAGGAACAAAAGAAACTGGAAAGGAAGATGC
TTTAGGAACATTTGCTTTAAATATATATTATAAGAGATCCTTCCATTATCT
GGGAAGATGTTTCAAACTCAGAGACAAGGACTTGGTTTGTAAGACAAACGATGATATGAA
GGCCTTGTAAAGAAAAAATAAAAGATGAAGTGTGAAA

FIGURE 16

MCSRVPLLLPLLLALGPGVQGCPSCQCSQPQTVCFTARQGTTVPRDVPPDTVGLYVFENGIT
MLDAGSFAGLPGLQLLDLSQNQIASLPSGVFQPLANLSNLDLTANRLHEITNETFRGLRRLERLY
LGKNRIRHIQPGAFDTLDRLLEKLQDNELRALPPLRPRLLLLDLSHNSLLALEPGILDGTANVE
ALRLLAGLGLQQLDEGLFSRLRNLDLVDSDNQLERVPPVIRGLRGLTRLRLAGNTRIAQLRPEDL
AGLAALQELDVSNLSQLPGDLSGLFPRLLLAAARNPFNCVCPLSWFGPWVRESHVTLASPEE
TRCHFPPKNAGRLLLELDYADFGCPATTTATVPTTRPVVREPTALSSSLAPTWLSPTAPATEAP
SPPSTAPPTVGPVPQPQDCPPSTCLNGGTCHLGTRHHLACLCPEGFTGLYCESQMGQGTRPSPTP
VTPRPPRSLTGLIEPVSPSTSRLVGLQRYLQGSSVQLRSRLTYRNLSGPDKRLVTLRLPASLAEY
TVTQLRPNATYSVCVMPLGPGRVPEGEEACEAHTPPAVHSNHAPVTQAREGNLPLLIAPALA
LLAALAAVGAAAYCVRRGRAMAAAQDKGQVPGAGPLELEGVKVPLEPGPKATEGGEALPSGSE
CEVPLMGFPGPGLQSPHLAKPYI

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

GCAGCGGGGAGGCCGGCGGTGGTGGCTGAGTCCTGGTGGCAGAGGCCAAGGCCACAGCTCATGCG
GGTCGGATAGGGCTGACGCTGCTGTGCGGTGCTGAGCTTGGCCTCGCGTCCTCGG
ATGAAGAAGGCAGCCAGGATGAATCCTAGATTCCAAGACTACTTGACATCAGATGAGTCAGTA
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTCTGATTCAAAGAAATCTGA
ATTAGAATCCTCTATTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGGAAAGTGTACAG
AAGATATCAGCTTCTAGAGTCTCAAATCCAGAAAACAAGGACTATGAAGAGCCAAGAAAAGTA
CGGAAACCAGCTTGCACGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTCT
TTTCCTAGATAAGGAGTATGATGAATGTACATCAGATGGAGGGAAAGATGGCAGACTGTGGTGTG
CTACAACTATGACTACAAAGCAGATGAAAAGTGGGGCTTTGTGAAACTGAAGAAGAGGCTGCT
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAACGTGAATGAAAATCTTAATGGAAG
CAATAAGAAAAGCCAAAAAGAGAAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA
CCAAAGCCCTGGAGAGAGTGTCAATGCTCTTATTGGTGTGATTACTGCCACAGAATATCAG
GCAGCGAGAGAGATGTTGAGAAGCTGACTGAGGAAGGCTCTCCAAAGGGACAGACTGCTTTGG
CTTTCTGTATGCCTCTGGACTTGGTGTAAATTCAAGTCAGGCCAAAGGCTCTGTATATTACAT
TTGGAGCTCTGGGGCAATCTAATAGCCCACATGGTTTGGTAAGTAGACTTTAGTGGAAAGGCT
AATAATATTAACATCAGAAGAATTGTGGTTATAGCGGCCACAACCTTTCAGCTTCATGATC
CAGATTGCTTGTATTAAGACCAATATTCAAGTTGAACCTCCTCAAATTCTGTTAATGGATAT
AACACATGGAATCTACATGTAATGAAAGTGGTGGAGTCCACAATTTCCTTAAATTTCTGAGTTG
TTGGCTGATTGCCCTAAAAAGAGAGATCTGATAAATGGCTCTTTAAATTTCTGAGTTG
GAATTGTCAGAATCATTTTACATTAGATTATCATAATTAAAATTTCTTGTAGTTCA
AAATTGGTAAATGGTGGCTATAGAAAAACACATGAAATATTACAAATATTGCAACAATGC
CCTAAGAATTGTTAAATTCATGGAGTTATTGTGCAGAATGACTCCAGAGAGCTACTTCTG
TTTTTACTTTCATGATTGGCTGCTTCCATTATTCTGGTCAATTGTTACTTTCTTGTAAATTGG
AAGATTAACTCATTTAATAAAATTATGCTAAGATTAaaaaaaaaaaaaaaaaaaaaaaa
aaa

100 90 80 70 60 50 40 30 20 10

FIGURE 18

MRVRIGLTLLLCAVLLSLASASSDEEGSQDESLSKTLTSDESVKDHTTAGRVVAGQIFLDSEEESEL
ESSIQEEEDSIKSQEGESVTEDISFLESPPNPKDYEEPKKVRKPALTAIEGTAHGEPCHPFLFLDK
EYDECTS DGDREDGRWCATTYDYKADEKWGFCETEEAAKRRQMQEAEMMYQTGMKILNGSNKKSQKR
EAYRYLQKAASMNHTKALERVSYALLFGDYLQPNIQAAREMFEKLTEEGSPKGQTALGFLYASGLGVN
SSQAKALVYYTFGALGGNLIAHMVLVSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCAAGATTTAAGCCCATTCTGCAGTGGAAATTCACTGAACACTAGCAAGAGGACACCACCTTCTT
GTATTATAACAAGAAAGGAGTGTACCTATCACACACAGGGGGAAAAATGCTCTTGTTGGGTGCTAGG
CCTCCTAACCTCTGTGGTTCTGTGGACTCGTAAGGAAAACTAAAGATTGAAGACATCACTG
ATAAGTACATTTTATCACTGGATGTGACTCGGGCTTGAAACTTGGCAGCCAGAACCTTTGAT
AAAAAGGGATTTCATGTAATCGCTGCCTGCTGACTGAATCAGGATCAACAGCTTAAAGGAGA
AACCTCAGAGAGACTTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGAATGTCAGAGGACTG
CCCAGTGGGTGAAGAACCAAGTGGGGAGAAAGGTCTCTGGGGTCTGATCAATAATGCTGGTGT
CCCGGCCGTGCTGGCTCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAA
CCTGTTGGACTCATCAGTGTGACACTAAATATGCTCCTTGGTCAAGAAAGCTCAAGGGAGAG
TTATTAATGTCTCCAGTGTGGAGGTGCCCTGCAATCGTGGAGGGCTATACTCCATCCAAA
TATGCAGTGGAAAGGTTCAATGACAGCTTAAGACGGGACATGAAAGCTTGGTGCACGTCTC
ATGCATTGAACCAGGATTGTC~~AA~~AAACTGGCAGATCCAGTAAAGGTAATTGAAAAAAAC
TCGCCATTGGGAGCAGCTGTC~~CC~~CAGACATCAAACAAATATGGAGAAGGTTACATTGAAAAAA
AGTCTAGACAAACTGAAAGGCAATAAAATCCATGTGAACATGGACCTCTCCGGTGGTAGAGTG
CAIGGACCACGCTCTAACAAAGTCTCTCCTAAGACTCATTATGCCGCTGGAAAGATGCCAAA
TTTCTGGATAACCTCTGTCTCACATGCCAGCAGCTTGCAAAGACTTTATTGTTGAAACAGAAA
GCAGAGCTGGCTAACCC~~CAAGG~~CAGTTGACTCAGCTAACCAACAAATGTCCTCCAGGCTATGA
AATTGCCGATTCAAGAACACATCTC~~TTTCAACCC~~ATTCTTATCTGCTCAACCTGGACT
CATTTAGATCGTGT~~TTTG~~GATTGCAAAAGGGAGTCCACCATCGCTGGTGTATCCCAGGGT
CCCTGCTCAAGTTCTTGGAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCCTGCC
GTATTTAGGCTTGCCTGCTGGTGT~~GATG~~TGAAGGGAAATTGAAAGACTTGCCCATTCAAAATGA
TCTTACCGTGGCCTGCCCATGCTTATGGCCCCAGCATTACAGTAACTTGTGAATGTTAAGT
ATCATCTTATCTAAATTTAAAGATAAGTCACCCAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAA

2020-09-20 10:45:00

FIGURE 20

MLFWVLGLLILCGFLWTRKGKLKIEDITDKYIFITGCDSFGNLAARTFDKKGFHVIACLTESG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVLAPTDWLTLLEDY
REPIEVNLFGLISVTLNMLPLVKKAQGRVINSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMK
AFGVHVSCIEPGLFKTNLADPVKIEKKLAIWEQLSPDIKQQYGEGYIEKSLDKLGNKSYVNMD
LSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQAKELANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGC GGCGGTAGCATGGAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTGTGCTCGGCG
CACTCGCTTCCAGCACCTAACACGGACTCGGACACGGAAGGTTCTTCTTGGGGAAAGTAAAA
GGTGAAGCCAAGAACAGCATTACTGATTCCAAATGGATGATGTTGAAGTTGTTATAACAATTGA
CATTCAAGAAATATATTCCATGCTATCAGCTTTAGCTTATAATTCTTCAGGCGAAGTAAATG
AGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAGAATGTGGTAGGTTGGTACAAATTCCGT
CGTCATTCA GATCAGATCATGACGTTAGAGAGAGGCTGCTCACAAAAC TTGCAGGAGCATT
TTCAAACCAAGACCTGTTCTGCTATTAACACCAAGTATAAACAGAAAGCTGCTCTACTC
ATCGACTGGAACATCCTTATATAAACCTCAAAAGGACTTTACAGGGTACCTTAGTGGTT
GCCAATCTGGGCATGCTGAACA ACTGGTTATAAAACTGTATCAGGTCCTGTATGTCCACTGG
TTTAGCCGAGCAGTACAAACACAGCTCAAATTGGTAAGAAGATGGATCCTAAAGGAGG
TACATAAGATAAAATGAAATGTATGCTTCATTACAAGAGGAATTAAAGAGTATATGCAAAAAGTG
GAAGACAGTGAACAAGCAGTAGATAAAACTAGTAAAGGATGTAACAGATAAAAGAGAAATTGA
GAAAAGGAGAGGAGCACAGATTCAAGCAGCAAGAGAGAAGAACATCCAAAAGACCTCAGGAGA
ACATTTCCTTGTCAAGCATTACGGACCTTTTCAAATTCTGAATTCTTCATTGTT
ATGCTTTAAAAAATAGACATGTTCTAAAGTAGCTGTAAC TACAACCACATCTCGATGTTAGT
AGACAATCTGACCTTAATGGTAGAACACACTGACATTCCCTGAAGCTAGTCCAGCTAGTACACCAC
AAATCATTAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTGGTTGTTA
GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAGCATCCAAAAT
GAGCAGCCCAGAACAGATGAAGAAAATTGAAAAGATGAAGGGTTTGGTAATATTACGGCTC
CTACATTTGATCTTTAACCTTACAAGGAGATTTTATTGGCTGATGGTAAAGCCAAAC
ATTCTATTGTTTACTATGTTGAGCTACTGCA GTAAAGTTCATTGTTTACTATGTTCA
TGTTGAGCTAATACAGATAACTCTTAGTGCATTACTTCACAAAGTACTTTCAAACATCA
GATGCTTTATTCCAAACCTTTTACCTTCACTAAGTGTGAGGGAAAGGCTTACACAG
ACACATTCTTGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCAGCACT
TAGGGAAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGCAACGTATT
GAGACCAGTCTATTAAAAAATGAAAAGCAAGAATAGCCTTATTGAAAATATGGAAA
GAAATTATGAAAATTATGAGTCATTAAATTCTCCTTAAGTGTACTTTAGAAGTA
CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAATTGCAAAACATCATCT
AAAATTAAAAAAAAAAAAAAAAAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTSDTEGFLLGEVKGEAKNSITDSQMDDVEVVYTIDIQKYI
PCYQLFSFYNSGEVNEQALKKILSNVKNNVGWYKFRRHSDQIMTFRERLLHKNLQEHSNQDL
VFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQAVDKLVKDVNRLKREIEKRGA
QTHSSKFFEDGSLKEVHKINEMYASLQEELKSICKVEDSEQAVDKLVKDVNRLKREIEKRGA
QIQAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHSVSKSSCNYNHHLDVVDNLTL
MVEHTDIPEASPASTPQIIKKALDLDWRQFKRSRLDTQDKRSKANTGSSNQDKASKMSSPET
DEEIEKMKGFGEYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGGCCGGAGGCAGAGTCAGCCGAGCAGTCAGCCGAGTCAGCCGACGAGCGGACCAGCGCAGGGCAGCCAA
GCAGCGCGCAGCGAACGCCGCCGCCACACCCCTCGCGTCCCCGGCGCCCTGCCACCCCTCCCTCCCC
GCCGTCCCCGCCCTCGCCGGCAGTCAGCTTGCGGGTTGCGTGCCTGCCGAAACCCCGAGGTCAACCAGCCGCCCT
GCTTCCCTGGGCCGCCGCCCTCCACGCCCTCTCTCCCTGCCCGCCTGGCACCGGGGACCGTTGCCTGA
CGCGAGGCCAGCTACTTTGCCCGCCTCTCCCGCTGCCTCGCCCTCTCCACCAACTCCAACCTCCTCTCCC
TCCAGCTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTGCCGCTGCCGTAGCGCCGCTCCCGTCCGGTCCAAA
GGTGGGAACCGCGTCCGCCCGGCCACCATGGCACGGTTGGCTTGCCCGCCTCTGCACCCTGGCAGTGCTC
AGCGCCGCGTGCCTGCCGAGCTCAAGTCAGAAAGTTGCTCGAAGTGCAGCTCTTACGTGCTCAAAGGCTTC
AACAAAGAACGATGCCCTCACAGAGATCAACGGTGTACATTGAAGATCTGCCCCAGGGTCTACCTGCTGCT
CAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGATGATTCAAAGTGTGGTCAGCGAACAGTGAATCATTG
CAAGCTGCTTGTACGTTACAGAAAGTTGATGAATTCTCAAAGAACTACTTGAAGATCTTGTGGTCAGCGAACAGTGAATCATTG
AATGATATGTTGTGAAGACATATGCCATTATACATGCAAAATTCTGAGCTATTAAAGATCTTGTAGAGTTG
AAACGTTACTACGTGGTGGGAAATGTGAACCTGGAAGAAATGCTAAATGACTCTGGCTGCCCTGGAGCGGATG
TTCCGCCCTGGTAACCTCCAGTACCACTTACAGATGAGTATCTGGAATGTGTGAGCAAGTACGGAGCAGCTGAAG
CCCTTCCGGAGATGCCCTGCCAAATTGAAAGTCCAGGTTACTCGTCTTTGTACAGCCCGTACTTCGCTCAAGGC
TTAGCGGTTGCCGGAGATGCTGAGCAAGGTCTCGTAAACCCACAGCCAGTGTACCCATGCCCTGTTGAAG
ATGATCTACTGCTCCACTGCCGGGCTCGTGAAGTGTGAAGCCATGTTACAACACTGCTCAAACATCATGAGAGGC
TGTTGGCAACCAAGGGATCTGATTGATGAAACAAATTCTAGATGCTATGCTGATGGTGCAGAGAGGCTA
GAGGGTCTTCAACATTGAATCGGTATGGATCCCATCGATGTGAAGATTCTGATGCTATTATGAACATGCAAGGAT
AAATGTTCAAGTGTCTCAGAAGGTTTCCAGGGATGTTGGACCCCCAAGCCCTCCAGTGGACGAATTCTCGT
TCCATCTCTGAAAGTCCCTCAGTGTGCTCAGACCACATCCCCGAGGAACGCCAACACAGCAGCTGGCACT
AGTTGGACCGACTGGTTACTGATGTCAGGAGAAACTGAAACAGGCAAGAAATTCTGGCTCCCTCCGAGCAAC
GTTTGCAACGATGAGAGGATGGCTGAGGAAACGGCAATGAGGATGACTGTTGAATGGAAAGGCAAAGCAGGTAC
CTGTTGCAGTGAACAGGAATGGATTAGCAACCCAGGGCAACAACCCAGGGTCCAGGTTGACACCAGCAAACAGAC
ATACTGATCTCGTCAAATCATGGCTCTCGAGTGTGATGACCAGCAAGATGAAGAATGCTACATAATGGAACGACGTG
GACTTCTTGTATCAGTGTGAAAGTAGTGGAGAAGGGTGAAGTGGAGCTGTGAGTATCAGCAGTGCCTCAGAG
TTTGACTACAATGCCACTGACCATGCTGGAAAGAGTGCAATGAGAAAGCCGACAGTGTGGTGTCCGTCTGGGCA
CAGGCCACCTCCTCAGTGTCTGCACTTGTCTCTGTTATGAGAGAGAGTGGAGATAATTCTCAAACCTGAG
AAAAAGTGTGATCAAAGTAAAGGACCAAGTATCACTTCTCTGTTATGAGAGAGAGTGGAGATAATTCTCAAACCTGAG
TGGACAAACATGTACAGTTTACTATGTGGCCACTGGTTAAGAAGTGTGACTTGTGTTCTCATTGAGTTGGG
AGGAAAAGGGACTGTCCTATTGAGTTGGTCTGCTCCCCAACCATGTTAAACGTGGCTAACAGTGTAGGTACAGAA
CTATAGTTAGTGTGATTTGTGATTTTACTCTATTATTTGTTGTGTTTTCTCATTCGTTGTGGGTT
TTTTTTCCAAGTGTGATCTGCCCTGTTCTACAAGCAAACCCAGGGTCCCTCTGGCACGTAACATGTACGTATT
TCTGAAATTTAAAGTGTACAGAACAGGTTTATTATGTTATCTTAAAGAAAAGCCAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAAELKSKCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQGST
CCSQEMEEKYSLQSKDDFKSVVSEQCNCNHQAVFASRYKKFDEFFKELLENAEKSNDMFVKTYGH
LYMQNSELFKDLFVELKRYVVGNVNLEEMLNDFWARLLERMFRILVNSQYHFTDEYLECVSKYTE
QLKPFGDVPRKLKLQVTRAFVAARTFAQGLAVAGDVSKVSVNPTAQCTHALLKMIYCSHCRL
VTVKPCYNCSNIMRGCLANQGDLDFEWNNFIDAMILVAERLEGPFNIESVMDPIDVKISDAIMN
MQDNSVQVSQKVFOQCGPPKPLPAGRISRSISESAFSARFRPHHPEERTTAAGTSLDRLVTDVK
EKLKQAKKFWSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYLFAVTGNGLANQGNNPEVQVDT
KPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAG
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glycans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTCCATTATATTCTTCAGCAACT
TACAGCTGCACCGACAGTTGCGATGAAAGTTCTAATCTCTTCCCTCCTGTTGCTGCCACTAA
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGTGCAGGAGGCCACAGGGACCGAGGC
CAGGCTTCTAGGAGATGGCTCAGGAAGGGGCCAAGAATGTGAGTGCAAAGATTGGTTCTGAG
AGCCCCGAGAAGAAAATTCATGACAGTGCTGGGCTGCCAAAGAAGCAGTGCCCTGTGATCATT
TCAAGGGCAATGTGAAGAAAACAAGACACCAAGGCACCCACAGAAAGCCAACAAGCATTCCAGA
GCCTGCCAGCAATTCTCAAACAATGTCAGCTAAGAAGCTTGCTCTGCCTTGTAGGAGCTCTG
AGCGCCCACCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAAGACACTC
TTCTTCTCCCACCTCACTCTCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAGCA
TGTTTTCAAGATCATTTGTTGCTCTCTAGTGTCTTCTCTCGTCAGTCTTAGCCT
GTGCCCTCCCTTACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAAACTGTAGCTCCT
AGCTAGTGTCAATTAAACCTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTT
AAATGTCAAAAAAAAAAAAAAAA

FIGURE 26

MKVLISSLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM
TVSGLPKKKQCPKCDHFKGNNVKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGCGCTGCAGAGGCTGAGCAGGGAAAAGCCAGTGCCCCAGCGGAAGCACAGCTCAG
AGCTGGTCTGCCATGGACATCCTGGTCCCCTCCTGCAGCTGCTGGTCTGCTTCTTACCTGCC
CCTGCACCTCATGGCTCTGCTGGCTGCTGGCAGCCCCGTGCAAAAGCTACTTCCCACCTGA
TGGCGTGTGACTCCAAGAGCAACCAGCAAGATGGAGAGCAAGAACGGAGCTTCAGCCAG
ATAAAGGGGTTACAGGAGCCTCCGGAAAGTGGCCCTACTGGAGCTGGCTGCGGAACCGGAGC
CAACTTCAAGTTCTACCCACCGGGCTGCAGGGTACCTGCCTAGACCCAAATCCCCACTTGAGA
AGTTCTGACAAAGAGCATGGCTGAGAACAGGCACCTCAATATGAGCGGTTGTGGTGGCTCCT
GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGCTGCACACTGGTGTG
CTCTGTGCAGAGCCAAGGAAGGTCCCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTG
TCTTTCTGGGAGCATGTGGCAGAACATGGAGCTGGCCTTCATGTGGCAGCAAGTTTC
GAGCCCACCTGGAAACACATTGGGATGGCTGCTGCCCTCACCAGAGAGACCTGGAAGGATCTGA
GAACGCCAGTTCTCCGAAATCAAATGAAACGACAGCCCCCTCCCTGAAGTGGTACCTGTTG
GGCCCCACATCATGGAAAGGCTGTCAAACAATCTTCCAAGCTCCAAGGCACTCATTTGCTCC
TTCCCCAGCCTCCAATTAGAACAGCCACCCACCAGCCTATCTATCTCCACTGAGAGGGACCTA
GCAGAATGAGAGAAGACATTCATGTACCACTACTAGTCCCTCTCCCCAACCTCTGCCAGGGC
AATCTCTAACTCAATCCCGCCTTCGACAGTAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG
AAACACTAGGACCCCTGTTGATCCTCAACTGCAAGTTCTGGACTAGTCTCCAACGTTGCCTC
CCAATGTTGTCCTTCCCTGTTGCTCCATGGTAAAGCTCCTCGCTTCCCTGAGGCTACAC
CCATGCGTCTCTAGGAACCTGGTCACAAAGTCATGGTGCCTGCATCCCTGCCAACGCCCCCTGAC
CCTCTCTCCCCACTACCACCTTCTGAGCTGGGGCACCAGGGAGAATCAGAGATGCTGGG
ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTGTCTCAAATATTTTTAATAATAGACGAA
ACCACG

DNA sequence analysis

FIGURE 28

MDILVPLLQLLVLLTLPLHIMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSQIKGL
TGASGKVALLELGCCTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDVVVCTLVLCVQSPRKVLQEVRVLRPGGVLFWEHVAEPYGSWAFMWQQVFEPTW
KHIGDGCLTRETWKDLENAQFSEIQMERQPPPLKWLPGPHIMGKAVKQSFPSKALICSFPSL
QLEQATHQPIYLPLRGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATTTGCCTATCCACCTCCCCAAGCCCCTTTACCTATGCTGCTGCTAACGCTGCTGCTGCT
GCTGCTGCTGCTAAAGGCTCATGCTGGAGTGGGACTGGTCGGTGCCAGAAAGTCTCTCTG
CCACTGACGCCCATCAGGGATTGGCCTCTTCCCCCTTCCTGTGTCTCCTGCCTCAT
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGAAGGGGAGAAAGTGGGGATGGCTA
GAAAAGCTGGGAGATAGGAACAGAAGAGGGTAGTGGTGGCTAGGGGGCTGCCTATTTAAA
GTGGTTGTTATGATTCTTATACTAATTATACAAAGATATTAAGGCCCTGTTCATTAAGAAATT
GTTCCCTCCCCGTGTTCAATGTTGTAAGATTGTTGTGTAATATGTCTTATAATAAAC
AGTTAAAAGCTGAAAAA~~AAAAAAAAAAAAAAA~~

FIGURE 30

MLLLTLLLLLKGSCLEWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHRPAMTC SQAQPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTTGAATTCTTCAACTATAACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAGTT
CCTCCAAGCAAGTCATTTCCCTTATTAAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT
ATTGCATCTGTTGATAAAATGATGTTGACACCCCTCCACCGAATTCTAAGTGGATCATGTCGG
GAAGAGATAACAATCCTTGGCTGTGTATCCTCGCATTAGCCTGCTTGGCATGATGTTACC
TTCAGATTCACTACCACCCCTCTGGTCACATTTCATTGTTATTGGGATTGTTGTT
TGTCTGCGGTGTTTATGGTGGCTGTATTATGACTATACCAACGACCTCAGCATAGAATTGGACA
CAGAAAGGAAAAATAIGAAGTGCCTGGGTTTGTCTATCGTATCCACAGGCATCACGGCAGTG
CTGCTCGTCTTGATTTGTTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTCCAAATCAC
AAATAAAGCCATCAGCAGTGCTCCCTCCTGCTGTTCCAGCCACTGTGGACATTGCCATCCTCA
TTTCTCTGGTCTCTGGTGGCTGTGCTGCTGAGCCTGGAACTGCAGGAGCTGCCAGGTT
ATGGAAGGCGCCAAGTGGATATAAGCCCTTCGGCATTGGTACATGTGGTGTACCTT
AATTGGCCTCATCGGACTAGTGAATTCACTCCTGCGTGCAGCAAATGACTATAGCTGGGCAG
TGGTTACTGTTATTCAACAGAAGTAAAAATGATCCTCCTGATCATCCCATCCTTCGTCTCTC
TCCATTCTCTTCTTACCATCAAGGAACCGTTGTGAAAGGGTCACTTAACTCTGTGGTGAG
GATTCCGAGAACATTGTATGTCATGCAACTGCAAAACGCACTGAAAGAACAGCAGCATGGTGCATTG
CCAGGTACCTGTTCCGATGCTACTGCTGTTCTGGTGTCTGACAAATACCTGCTCCATCTC
AACCGAGATGCATATACTACAACGTCTATTAAATGGGACAGATTCTGTACATCAGCAAAGATGC
ATTCAAATCTTGTCCAAGAAACTCAAGTCACCTTACATCTATTAACTGCTTGGAGACTTCATAA
TTTTCTAGGAAAGGTGTTAGTGGTGTGTTCACTGTTGGAGACTCATGGCTTTAACTAC
AATCGGGCATTCCAGGTGGCAGTCCCTCTGTTATTGGTAGCTTTGCCTACTTAGTC
CCATAGTTTTATCTGTGTTGAAACTGTGCTGGATGCACCTTCTGTGTTGCTGTTGATC
TGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTATGGATCAAGAATTCTGAGTTCGTA
AAAAGGAGCAACAAATTAAACAAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGAATGAGGA
GGGAACAGAACCTCAGGCCATTGTGAGATGATACCCATTAGGTATCTGTACCTGGAAAACATT
TCCTCTAAGAGCCATTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATT
TTAAAAGACCTAATAACCCATTCTCTCAAAA

FIGURE 32

MSGRTILGLCILALALSLAMMFTFRFITTLLVHIFISLVLGLLFVCGVLWWLYYDYNNDLSIE
LDTERENMKCVLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQPLWTFA
ILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQQMTIA
GAVVTCYFNRSKNDDPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVMMQNALKEQQHG
ALSRYLFRCYCFCFWCLDKYLLHLNQNAYTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFGD
IIIFLGKVLVVCFITVFGGLMAFNYNRAFQVWAVPLLVAFFAYLVAHSFLSVFETVLDALFLCFA
VDLETNDGSSEKPYFMDQEFLSFVCRSNKLNNARAQQDKHSLRNEEGTELQAIVR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33

GTTCGATTAGCTCCTCTGAGAAGAGAAAAGGTTCTGGACCTCTCCCTGTTCTCCCTAGA
ATAATTGTATGGGATTGTGATGCAGGAAGCCTAACGGAAAAAGAATATTCAATTCTGTGTGGT
GAAAATTTTGAAAAAAATGCCCTCTCAAACAAAGGGTGTCAATTCTGATATTTGAGGAC
TGTTGTCTCACTATGAAGGCATCTGTTATGAAATGTCCCTGTTGTGACTGGAGTAC
ATTCAAACAAAGAACGCGAAAGAAGATTAAAAGGCCAAGTCAGTGCCTCAGATCAACTGC
GATGTCAAAGCGGAAAGATCATCGATCCTGAGTCATGTGAAATGTCCAGCAGGATGCCAAGA
CCCCAAATACCATGTTATGGCACTGACGTGATGCATCCTACTCCAGTGTGTGGCGTGCCG
TACACAGTGGTGTGCTGATAATTCAAGGAGGGAAAATACTTGTTGCGAAGGTTGCTGGACAGTCT
GGTTACAAAGGGAGTTATTCCAACGGTGTCCAATGTTATCCCTACCACGATGGAGAGAACCTT
TATCGTCTTAGAAAGTAAACCCAAAAGGGTGTAACTAACCATCACGCTTACATACATCAT
CGAAAAGTCAGCTGCCAAGCAGGTGAGACCACAAAGCCTATCAGAGGCCACCTATTCCAGGG
ACAAC TGCA CAGCGG TCA CTGATGCAGCTCTGGCTGTCACTGAGCTGTGGCCACCCCCAC
CACCTTGCCAAGGCCATCCCCTCTGCTGCTTCTACCACCA GATCCCCAGACCAAAATCAGTGG
GCCACAGGAGCAGAGATGGATCTCTGGTCCACTGCCACCTACACAAGCAGCCAAAAGGCC
AGAGCTGATCAGGTTATCCAAAGGCAAGATCCTTCAGGAGCTGCCTCCAGAAACCTGTTGGAGC
GGATGTCAGCTGGGACTGTTCCAAAAGAAGAATTGAGCACACAGCTTTGGAGCCAGTATCCC
TGGGAGATCAGTCAAAATTGACTTGTGCTTTTAATTGATGGGAGCACAGCATGGCAA
CGGGCATTCCGAATCCAGAAGCAGCTCTGGCTGATGTTGCCAAGCTCTGACATTGGCCCTGC
CGGTCCACTGATGGGTGTTGTCAGTATGGAGACAACCCCTGCTACTCACTTAACCTCAAGACAC
ACACGAATTCTCGAGATCTGAAGACAGCCATAGAGAAAATTACTCAGAGAGGAGGACTTTCTAAT
GTAGGTCGGGCCATCTCTTGTGACCAAGAACCTCTTCCAAAGCCAATGGAAACAGAACGG
GGCTCCAATGTGGGGTGTGATGGTGGATGGCTGGCCACGGACAAAGTGGAGGAGGCTCAA
GACTTGCAGAGAGTCAGGAATCAACATTCTTCATCACCAATTGAAAGGTGCTGCTGAAAATGAG
AAGCAGTATGTGGTGGAGCCCAACTTGTCAAACAAAGGCCGTGTCACAACAAACGGCTCTACTC
GCTCCACGTGCAGAGCTGGTTGGCCTCCACAAGACCCCTGCAGCCTCTGGTGAAGCGGGTCTGCG
ACACTGACCGCCTGGCTGCAGCAAGACCTGCTGAACCTGGCTGACATTGGCTCATCGAC
GGCTCCACGAGTGTGGGACGGCAACTTCCGCACCGTCTCCAGTTGTGACCAACCTCACCAA
AGAGTTGAGATTCCGACACGGACACGCGCATCGGGCGTGCAGTACACCTACGAACAGCGGC
TGGAGTTGGGTTGACAAAGTACAGCAGCAAGCCTGACATCCTCAACGCCATCAAGAGGGTGGGC
TACTGGAGTGGTGGCACCAGCACGGGCTGCCATCAACTTGCCTGGAGCAGCTTCAAGAA
GTCCAAGCCCAACAAAGAGGAAGTTAATGATCCTCATCACCGACGGAGGTCTACGACAGTCC
GGATCCCGCCATGGCTGCCCATCTGAAGGGAGTGTACACCTATGCGATAGGCGTTGCCTGGCT
GCCCAAGAGGAGCTAGAAGTCATTGCCACTCACCCGCCAGAGACCACTCCTCTTTGTGGACGA
GTTTGACAAACCTCATCAGTATGTCCCCAGGATCATCCAGAACATTGATCAGAGTTCAACTCAC
AGCCTCGGAACTGAATTCAAGACCAAGGAGCAGACCAAGCAAGTGTGCTTTACTAACTGACGTGTT
GGACCAACCCACCGCTTAATGGGGCACGCGCACGGTGCATCAAGTCTGGGCAGGGCATGGAGAAC
AAATGTTGTTATTATTCTTGCCTCATGCTTTCATATTCCAAAACCTGGAGTTACAAAGA
TGATCACAAACGTATAGAATGAGCAAAGGCTACATCATGTTGAGGGTGTGAGGATTTACAT
TTTGACAAATTGTTCAAAATAATGTTGGAATACAGTGCAGCCCTACGACAGGCTTACGTAG
AGCTTTGTGAGATTGTTAGTTAAGTGTATTCTGATTTGAACTCTGTAACCCCTCAGCAAGTTCT
TTTGTCATGACAATGTAGGAATTGCTGAATTAAATGTTAGAAGGATGAAAATAAAAAAAAAA
AAG

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG
CQDPKYHVVYGTDVASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSLSLPRWR
ESFIVLESKPKKGVTVPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQLLAVTVAVA
TPTTLPRPSPSAASTTSIPRQPQSVGHRSQEMDLWSTATYTSSQNRPRADPGIQQRQDPSGAAFQKP
VGADVSLGLVLPKEELSTQSLEPVSLGDPNCIDLSFLIDGSTSIGKRRFRIQKQOLLADVAQALDI
GPAGPLMGVVQYGDNPATHFNLKHTNSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPNVVVVMVWDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKTQLQPLVKRVCDTDRLACSKTCLNSADIGFVIDGSSSGTGNFRTVLQFVTN
LTKEFEISDTDRIGAVQYTYEQRLEFGFDKYSSKPDILNAIKRVGWWSGGTSTGAIAINFALQL
FKKSCKPNKRKLMILITDGRSYDDVRIPAMA AHLKGVITYAIGVAWAQEELEVIATHPARDHSFF
VDEFDNLHQYVPRIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCGTTAGGAGGGCTGCGTGTGGAAAAGCTATCAAGGAAGAAATTGC
CAAACCATGTTTTCTGTTTCAGAGTAGTCACAACAGATCTGAGTGTAAATTAGCATGGAAT
ACAGAAAACAACAAAAACTTAAGCTTAATTCATCTGGATTCCACAGTTCTTAGCTCCCTGGACCC
GGTTGACCTGTTGGCTCTCCCGCTGGCTGCTATCACGGTGTCTCCGACTACTCACCCGAGTGTA
AAGAACCTCGGCTCGCGTCTGAGCTGCTGTGGATGGCCTCGGCTCTCTGGACTGCTCCGAGTA
GGATGTCACTGAGATCCCTAAATGGAGCCTCGCTGCTGCACTCCTGAGTTCTTGATGTGGTAC
CTCAGCCTCCCCACTACAAATGTGATAGAACCGGTAACTGGATGTAATTCTATGAGTATGAGCGATTAA
CAGACAAGACTTCACCTCACACTCGAGAGCATTCAAACACTGCTCTCATCAAATCCATTCTGGTCATT
TGGTGACCTCCCACCCCTCAGATGTGAAAGCCAGGCCATTAGAGTTACTTGGGGTGAAGAAAAGTCT
TGGTGGGGATATGAGGTTCTACATTTCATTAGCCAAGAGGCTGAAGAGAACAAATGTTGGC
ATTGTCCTAGAGGATGAACACCTCTTATGGTACATAATCCGACAAGATTAGACACATATAATA
ACCTGACCTGAAAACCATTATGGCATTAGGTGGTAACTGAGTTTGCCCCAATGCCAAGTACGTAATG
AAGACAGACACTGATGTTCTACAAACTGGCAATTAGTGAAGTATCTTAAACCTAAACACTCAGA
GAAGTTTCACAGGTTATCCTCTAATTGATAATTTCCTATAGAGGATTACCAAAACCCATATT
CTTACCAAGGAGTATCCTTCAGGTGTTCCCTCCTACTGCAGTGGGTTGGTTATATAATGTCAGAGAT
TTGGTGCCAAGGATCTATGAAATGATGGTCAGTAAACCCATCAAGTTGAAGATGTTATGTCGGGAT
CTGTTGAATTATTAAAAGTGAACATTCAATTCCAGAAGACACAATCTTCTTCTATAGAATCC
ATTGGATGTCGCAACTGAGACGTGTATTGCAGCCATGGCTTCTTCAAGGAGATCATCACTTT
TGGCAGGTATGCTAAGAACACACATGCCATTATTAACTCACATTCTACAAAAGCCTAGAAGGACAG
GATACCTTGGAAGTGTAAATAAAAGTAGGTACTGTGGAAAATTCACTGGGAGGTCACTGTGCTGGCTT
ACACTGAAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTGTGATTATTAGTCAGG
CCCTCAAAAGATGATATGGAGGAATTAAATATAAGGAATTGGAGGTTTGCTAAAGAAATTAGG
ACCAAAACAATTGGACATGTCATTCTGTAGACTAGAATTCTAAAGGGTGTACTGAGTTATAAGCTCA
CTAGGCTGAAAAACAAACATGTAGAGTTTATTGAACAAATGTAGTCATTGAAAGGTTTGTA
TATCTTATGTGGATTACCAATTAAAAATATGTAGTTCTGTGTCAAAAGCTTCAACTGAAAGTT
CTGAACAAAATTTCACCTGTTGGCATTAAAGTACTTCAGTCAAGATGTTGAGTATTCACTGAA
ATTATTAAATTACTCAACTTGTGTTAAATGTTGACGATTCAACAAAGATAAAAGGATAG
TGAATCATTCTTACATGCAAACATTCCAGTTACTTAACGTAGTCACTGTTATTGATA
TTAATGAAAGTCAGGTCAATTGCAATCAGTAATCTCTGGACTTGTAAATATTACTGTGGT
AATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPIYRQDFHF
TLREHSNCSSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSSWWGYEVLTFFLLGQEAEKDMLA
LSLEDEHLLYGDIIRQDFLDTYNNLTKTMAFRWVTEFCPNAKYVMKTDTDVFINTGNLVKYLL
NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCSGLGYIMSRDLVPRIYEMMGHV
KPIKFEDVYVGICLNLLKVNIHIPEDTNLFFLYRIHLDVCQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

CGCTCGGGCACCAAGCAGGGCAAGGATGGAGCTGGGTTGCTGGACGCCAGTTGGGGCTCACTTTCTTCAGCTCCTTCATC
TCGTCCCTGCCAAGAGAGTACACAGTCATTAAATGAAGCCTGCCCTGGAGCAGAGTGGAAATATCATGTGTCGGAGTGCTGTG
AATATGATCAGATTGAGTGCCTGCTGCCCGGAAAGAGGGAAAGTCGTGGTTATACCATCCCTGCTGCAGGAATGAGGAGAA
TGAGTGTGACTCCTGCCCTGATCCACCCAGGGTGTACCATCTTGAAAATCTGCAAGAGCTGCCGAAATGGCTCATGGGGGGT
ACCTTGGATGACTTCTATGTGAAGGGGTTCTACTGTGCAGAGTGCCTGGAGCAGGCTGTACGGAGGAGACTGCATGCGATGTG
GCCAGGTTCTGCCGAGCCCCAAAGGGTCAGATTTGTGGAAAGCTATCCCCTAAATGCTCACTGTGAATGGACCATTCACTGC
TAAACCTGGGTTGTCACTCCAACTAAGATTGTACATGTGAGTCTGGAGTTGACTACATGTGCCAGTATGACTATGTGAG
GTTCGTGATGGAGACAACCCGATGCCAGATCATCAPCGTCTCTGGCACPGAGCGGGCAGCTCCATCCAGACATAG
GATCCTCACTCCACGTCTCTCCACTCCGATGGCTCCAAGAATTGACGGTTCCATGCCATTATGAGGAGATCACAGC
ATGCTCTCATCCCTGTTCCATGACGGCACGTGCTGCCCTGACAAGGCTGGATCTACAGTGTGCCTGCTGGCAGGC
TATACTGGGCAGCCTGTGAAAATCTCCCTGAAAGAAAGAAACTGCTCAGACCCCTGGGGCCAGTCATGGTACCAAG
TAACAGGGGCCCTGGCTTATCACCGGACGCCATGCTAAATTGGCACCGTGGTGTCTTCTTTGTAACAACCTCATG
TCTTAGTGGCAATGAGAAAAGAACCTGGCAGCAGAATGGAGAGTGGTCAGGGAAACAGCCCATCTGCATAAAAGCCTGCCGA
GAACCAAAGATTTCAGACCTGGTGAGAAGGGAGAGTCTCCGATGCAAGGGAGACACCATTACACCAGCTAT
ACTCAGGGCCCTCAGCAAGCAGAAACTGCAAGTGCCTCCACTAACAGAACAGCCCTCCCTTGGAGATCTGCCATGG
ATACCAACATCTGCATACCCAGCTCCAGTATGAGTGCATCTACCCCTCTACCGGCCCTGGCAGCAGGGAGACATGT
CTGAGGACTGGGAAGTGGAGTGGCGGGCACCATCTGCACTCCCTATCTGCGGGAAATTGAGAACATCACTGCTCAAAGA
CCCAAGGGTTGCGCTGGCGTGGCAGGCACCATCTACAGGAGGACAGCGGGGTGCATGACGGCAGCCTACACAAGGGAGC
GTGGTTCTAGTGCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGTGGCTGCCACTGTGTTACTGACCTGGGAAG
GTCACCATGATCAAGACAGCAGACCTGAAAGATTGTTGGGAAATTCTACCGGGATGATGACCGGGATGAGAACACCAC
AGACCCCTACAGATTTCTGCTATCATTCTGCATCCAACTATGACCCCATCCTGCTGACATGCCATCCTGAAGCT
CCTAGACAAGGCCGTATCAGCACCCAGTCCAGCCACTCTGCTCGCTGCCAGTGGGATCTCAGCAGTCCCTCAGGAG
TCCACACATCTGTTGGCTGGAAATGCTCTGGCAGACGTGAGGAGGCCCTGGCTCAAGAACACACTGCGCTCTGGGG
TGGTCAGTGTGGTGGACTCGCTGCTGTGAGGAGCAGCATGAGGACATGGCATCCAGTGAGTGTCACTGATAACATGTT
CTGTGCCAGCTGGAACCCACTGCCCTCTGATATCTGCACTGCAGAGACAGGAGGACATGCCGCTGTGCTTCCCGGG
CGACCATCTGCCAGCCAGCTGCACTGGACTGGTCAGCTGGAGCTATGATAAAACATGCAAGGCCACAGGCTCTCCA
CTGCCCTCACCAAGGTGCTGCCCTTTAAAGACTGGATTGAAAGAAATGAAACCATGCTCATGCACTCCTGAGAAC
TGTTCTGATATCCGCTGTACGTGTCATTGCGTGAAGCACTGTGAGTGTGATTTGGCTGTGAACCTGGCT
GTGCCAGGGCTCTGACTTCAGGGACAAAACACTCACTGAAAGGGTGAAGTACCTCCATTGCTGGTAGGCTGATGCCGCTCCA
CTACTGGACAGCCAATTGGAAGATGCCAGGGCTTGCAGAACAGTAAGTTCTCAAGAACAGGACATATACAAAACCTCTCCA
CTCCACTGACCTGGGGCTTCCCAACTTCTGAGTATACGAATGCCATCAGCTGACCCAGGGAGCAGAGCTGGGATGTGGTGCATGCC
GCCCTTGTGAGGCTCTCAAGTCTAGAGAGCTGCCCTGGGACAGCCAGGGCAGCAGAGCTGGGATGTGGTGCATGCC
TGTGTACATGCCACAGTACAGTCTGGCTTCTCCCTCCATCTTGTACACATTAAATAAGGGTTGGCT
GAACATACAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAA

FIGURE 38

MELGCWTQLGLTFLQLLISLIPREYTVINEACPGAEWNIMCRECCYEYDQIECVC PGKREVVGYT
IPCCRNEENECDSCLIH PGCTIFENCKSCRNGSWGGTLDDFYVKGFYCAECRAGWYGGDCMRCGQ
VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDGQII
KRCVCGNERPAPIQSIGSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC
ACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNNSYVLSGNE
KRTCQQNGEWSGKQPICIKACREP KISDLVRRRVLPMQVSRETPLHQLYSAAFSKQKLQSAPTK
KPALPF GDLPMGYQHLHTQLQYECISP FYRRLGSSRTCLRTGKWSGRAPSCI PICGK IENITAP
KTQGLRWPWQAAIYRR TSGVHDGSLHKGAFLVCS GALVNERTVVVAAHCVTDLGKV TMK TADL
KVVLGKFYRDDDRDEKTIQSLQISAIILHPNYDPILL DADIAILKLLDKARISTRVQPI CLAAS R
DLSTS FQESHITVAGWNVLADVRSPGFKN DTLRGVVS VVD SLLCEQHEDHGIPVS VTDNMFC A
SWEPTAPS DICTAETGGIAAVSF PGRASPEPRWHLMGLVWSYDKTCSHRLSTAFTKVL PFKDW I
ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTTCCCTACATCCTCTCATCTGAGAACATCAGAGAGCATAATCTTACGGGCCGTGATTATTAAACGTGGCTTAATC
TGAAGGTTCTCAGTCAAATTCTTGATCTACTGATTGTGGGGCATGGCAAGGTTGCTAAAGGAGCTGGCTGG
TTGGGCCCTGTAGCTGACAGAAGGTGGCAGGGAGAATGCAGCACACTGCTCGAGAATGAAGGCGCTTCTGTGC
TGGTCTTGCCTGGCTCAGCTCTGCTAACTACATTGACAATGTGGCAACCTGCACCTCCTGTATTCAA
AAGGTGCCCTCCACTACGGCCTGACCAAAAGATAGGAAGAGGCCTGACAAGATGGCTGTCCAGACGGCTGTGCAGGCC
TCACAGCCACGGCTCCCCTCCCAGAGGTTCTGCAGCTGCCACCATCCTTAATGACAGACGACGCCAGCAG
ACCCGCCTACGTGTCCCTGGCAGAGGACGGCAGGCCAGCAATCAGCCAGTGGACTCTGGCCGGAGCAACCGA
GGGCACGGCCCTTGAGAGATCCACTATTAGAACGAGATCATTAAAAAAATTCGAGCTTGTAGTGTCTCGAA
GGACAAAGAGCGGGAGTGCAGTTGCAACCAGCCGACAGGGCAGGGAAATTCTGAAAACACCACTGCCCTGAAG
TCTTCCAAGGTTGTACACCTGATTCCAGATGGTAAATTACAGCATCAAGATCAATCGAGTAGATCCCAGTGA
GCCCTCTATTAGGCTGGTGGAGGTAGCAGAACCCACTGGCCATATCATTATCAACACATTATCGTATGGGG
TGATGCCAGAGACGCCGGCTACTGCCAGGAGACATCATTCTAAAGGTCACCGGATGGACATCAGCAATGCC
ACAACATACGCTGTGCCTCCTGCCAGCCCTGCCAGGTGCTGGCTGACTGTGATGCGTGAACAGAAGTCC
GCAGGAAACATGGACAGGCCGGATGCCATACAGACCCCGAGATGACAGCTTCATGTGATTCTCAACAAAAGTAG
CCGAGGGCAGCTGGAAATAAAACTGGTGGCAAGGTGGATGAGCCTGGGTTTCATCTCAATGTGCTGGATGGCG
GTGTGGCATATCGACATGGTCAGCTTGAGGAGAATGACCGTGTGTTAGCCATCAATGGACATGATCTCG
GCCAGAAAAGTGGCAGCTCATCTGATTCAAGGCCAGTGAAGAGCGTGTGACCTCGTGTGCCCCAGGTTGGCAGC
GGAGCCCTGACATCTTCAGGAAGCCGGCTGGAAACAGCAATGGCAGCTGGTCCCCAGGGCAGGGAGAGGACA
CTCCCAAGCCCTCCATCTACAAATTACTGTGATGAGAACGTTGGTAAATATC
GCATGACCGTCAGGGGAGCATCACATAGAGAACGTTGGTAAATATC
GCATGACCGTCAGGGGAGCATCACATAGAGAACGTTGGTAAATATC
TCATAAGCAGAGATGGAGAAATAAAACAGGTGACATTGGTAAATGGT
GGAGTGGCAGTGGCATTATTGAAAGAACATCATCCTCGATAGTACT
CCCAGGAAGACTGCAGCAGCCAGCAGCCCTGGACTCCAACCAC
GGGTGATGTGGATTACACGGTGCTTGATAACTGTAAAGATATTG
TGGGTTCTGCATTGTAGGAGGTTATGAAGAACATGGAAACAC
CACCAGCATAATGAGGAAATTAGATGTGGTAAATTCTCTG
TACATGCTTGCTGGCAAGACTGCTGAAAGAACTTAAAGGA
TTTATAGAATCAATGATGGTCAGAGGAAACAGAAAAT
TTGTCAGTTTATATTAAAGAAAGAACATATTG
ACACCTCAGAAAATATGATCC
AACATTGTTATATTCTATTCAATA
CAAGCTGATT
AAATTTTAAATG
App_ID=10063670

FIGURE 40

MKALLLVL PWLSP ANY IDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDCASLTATAPS
PEVSAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDGRSNRTRARPFERSTIRSRSFKKINR
ALSVLRRTKSGSAVANHADQGRENTTAAPEVFPRLYHLIPDGEITSIKINRVDPSESLSIRLV
GGSETPLVIIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLLRQPCQVLWLTVM
REQKFRSRNNGQAPDAYPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGVFIFNVLDGGVAYRHG
QLEENDRVLAINGHDLRYGSPEAAHLIQASERRVHLVVSQRSPDIFQEAGWNSNGSWSPG
PGERSNTPKPLHPTITCHEKVVNIIQKDPGESLGMTVAGGASHREWDLPIYVISVEPGGVISRDGR
IKTDILLNVDGVELTEVRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPP
SDWSPSWVMWLELPRLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG
RIRC DILLAVNGRSTSGMIHA CLARLLKELKGRI TLIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCAAGGCATTGTATCTCAGTTGTCATCAAGTCCAACTCAGATTGGAAAAGCTCAACTTGAAGCTTT
CTTGCCTGCAGTGAAGCAGAGAGATAATTATTACCGTAATAAACATGGGCTTCAACCTGACT
TTCCACCTTCCTACAAATTCCGATTACTGTTGCTGTTGACTTGTGCTGACAGTGGTTGGGTGGGC
CACCAAGTAACACTTCGTTGGGCCATTCAAGAGATTCTAAAGCAAAGGAGTTCATGGCTAATTCC
ATAAGACCCCTCATTGGGGAAAGGAAAAACTCTGACTAATGAAGCATCCACGAAGAAGGTAGAACTT
GACAACGTCCCTCTGTGTCCTTACCTCAGAGGCCAGAGCAAGCTCATTCAAACCAGATCTCAC
TTTGGAAAGAGGTACAGGCAGAAAATCCAAAGTGTCCAGAGGCCGGTATGCCCTCAGGAATGTAAG
CTTTACAGAGGGTCGCCATCCTCGTTCCCCACCGGAACAGAGAGAACACCTGATGTACCTGCTGGAA
CATCTGCATCCCTCCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTATCCACCAGGCTGAAGG
TAAAAAGTTAATCGAGCCAAACTCTGAATGTGGGCTATCTAGAACGCTCAAGGAAGAAAATTGGG
ACTGCTTATATTCCACGATGTGGACCTGGTACCCGAGAAATGACTTTAACCTTACAAGTGTGAGGAG
CATCCAAAGCATCTGGTGGTGGCAGGAACAGCACTGGTACAGGTTACGTTACAGTGGATATTGG
GGGTGTTACTGCCCTAACGAGAGCAGTTTCAAGGTGAATGGATTCTCTAACAAACTACTGGGAT
GGGGAGGCGAAGACGGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTCCGGCCCTG
CCTGAAGTGGTAAATAACATGGTCTTACACTAGAGACAAAGGAATGAGGTGAACGCAGAACG
GATGAAGCTCTTACACCAAGTGTACGAGTCAGGAGTGGGTTGAGTAGTTGTTCTTAAAT
TAGTATCTGTGGAACACAATCCTTATATATCACACATCACAGTGGATTCTGGTTGGTGCATTGACCC
TGGATCTTGGTGAATGTTGGAAGAAACTGATTCTTGTTGCAATAATTGGCTAGAGACTCAA
ATAGTAGCACACATTAAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTCCTTGTATTCT
TAGCAGAGCTCCTGGTGAATGAGTATAAAACAGTTGTAACAAGACAGCTTCTAGTCATTGAT
CATGAGGGTAAATAATGTAATATGGATACTGAAAGGACTTATATAAAAGGATGACTCAAAGGATAA
AATGAACGCTATTGAGGACTCTGGTGAAGGAGATTATTAAATTGAAAGTAATATATTGGGAT
AAAAGGCCACAGGAAATAAGACTGCTGAATGTCAGAGAACCCAGAGTTGTTCTCGTCCAAGGTAGAA
AGGTACGAAGATAACAATCTGTTATTCACTTCTGTACAATCATCTGTGAAGTGGTGGTGTCAAGGT
GAGAAGGCAGTCCACAAAGAGGGAGAAAGGCAGCAATCAGGACACAGTGAACCTGGGAATGAAGA
GGTAGCAGGAGGGTGGAGTGTGGCTGCAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGATAGC
CTTCAGGGGAGGACTGCCAGGTATGCCTCCAGTGAATGCCCACCAAGAGAATACATTCTTATTAGT
TTTAAAGAGTTTGAAATGATTTGACAAGTAGGATATGAATTAGCAGTTACAAGTTACAT
ATTAACATAATAATATGTCTATCAAATACCTCTGTAGTAAATGTGAAAAGCAAA

FIGURE 42

MGFNLTFHLSYKFRLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLILGKGKTLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQOLDYGIYVIHQAEKKFNRAKLLNVGYLEALKEENWDCFIIFHDV
DLVPENDFNLYKCEEHPKHLVGRNSTGYRLRSGYFGGTALSREQFFKVNNGFSNNYWGWGGED
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGGATGGCACTGAGCTCCAGATCTGGG
CCGCTTGCTCCTGCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTCCCACAAACAG
ACGGGACAACCTGCAGAGCTGCAACCCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCCAT
GTTCCAGAGGCGAAGGAGGCAGACACCCACTTCCCCATCTGCATTTCCTGCTGCGGCTGCTGTC
ATCGATCAAAGTGTGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTGCCCGTCCCCTCCC
TTCCTTATTATTCTGCTGCCAGAACATAGGTCTTGAATAAAATGGCTGGTTCTTTGTTT
TCCAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 44

MALSSQIWAACLLLLL~~LL~~ASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTHFPI
CIFCCGCCHRSKCGMCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCCAACATGCCTCACCC
TCATCTATATCCTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCGTGAAAGAGCTGGTCGGT
TCCGTTGGTGGGGCCGTGACTTCCCCCTGAAGTCAAAGTAAGCAAGTTGACTCTATTGTCTG
GACCTTCACACAAACCCCTCTGTCACCATACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAACGTGAAG
AAGAATGACTCAGGGATCTACTATGTGGGGATATAACAGCTCATCACTCCAGCAGCCCTCCACCCA
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGAAGAGGATGTGATT
TATACCTGGAAGGCCCTGGGCAAGCAGCCAAATGAGTCCCATAATGGGTCCATCCTCCCCATCTC
CTGGAGATGGGAGAAAGTGAATGACCTTCATCTCGGTTGCCAGGAACCCGTCAAGCAAACCT
TCTCAAGCCCCATCCTGCCAGGAAGCTCTGTGAAGGTGCTGATGACCCAGATTCCCTCCATG
GTCCTCCTGTGTCTCTGTTGGTCCCCCTCTGCTCAGTCTTTGACTGGGCTATTTCTTG
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGTGGACATTGTCGGG
AAACTCCTAACATATGCCCTGATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
AGAACAAATCCTAAAGGAAGATCCAGCAAATACGGTTACTCCACTGTGGAAATACCGAAAAAGAT
GGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTGCCTATGAGAATGTTA
TCTAGACAGCAGTGCACTCCCCTAAGTCTGCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSVGGAVTFPLSKVKQVDSIVWTFNTTPLVTIQP
EGGTIIITQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLTCMEHGEEDVIYTWKALGQAANESHNGSILPISWRWGESDMTFIC
VARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLIVPLLLSSLFVLGLFLWFLKRERQEYIE
EKKRVDICRETPNICPHSGENTEYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTCTGAGCCAGGGGTGACCAATGACCTGCTGCGAAGGATGGACATCCTGCAATGG
ATTCAGCCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCCAATGCGATAACCTCTAATTGTCAGCT
TAGTTGAGGAAGACCAATTTCTAAAACCCCATCTCTTGCTTGAGTGGTGGTCCCAGGAATT
ATAGGAGCAGGTCTGATGCCATTCCAGCAACAACAATGCTTGACAGCAAGAAAAAGAGCGTG
CTGCAACAACAGAACTGGAATGTTCTTCATCATTTTCAGTGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAGGTCTCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCAATTGTGAATTTCATTGAAAAACATCAGTGACATTCATCAGAACATCCTCAA
CTTGCAGTGGTTTCCAATGACTCTTGTGCACCTCCTACTGGTTCAATAAAACCCACCAGTAACG
ACACCAATGGCGAGTGGCTGGAGAGCATCTAGTTCCACTTCGATTCTGAAGAAAACAAACATAGG
CTTATCCACTCTCAGTATTAGGTCTATTGCTTGGAAATTCTGGAGGTCTGTTGGC
CAGTCAGATAGTCATCGGTTCCCTGGCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG
TGTAGTTAATGGGAATAAAATGTAAGTATCAGTAGTAGTTGAAAAA

FIGURE 48

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGLIGAGLMAIPA
TTMSLTARKRACCNNRTGMFLSSFFSVITVIGALYCMILISIQALLKGPLMCNSPSNSNANCEFSL
KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRLIHFSVFLGL
LLVGILEVLFGLSQIVIGFLGCLCGVSKRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCTGCCAGCTCAGGTGAGCCCTGCCAAGGTGACCTCGCAGGACACTGGTGA
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
ATCCGTGGGCTGCAGACCCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCTCGAACTGTGA
CATGGAGAGAGTGCACCTGGCCCTCTCCTACTGGCAGGCCTGACTGCCTTGGAAAGCCAATGACC
CATTTGCCAATAAACAGATCCCTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
TGCAGGGCTCCTGGCATTGCTGGGATCGCGCAGTTCTGAGTGGCAAATGCAAATACAAGAG
CAGCCAGAACAGCACAGTCCTGTACCTGAGAAGGCCATCCACTCATCACTCCAGGCTCTGCCA
CTACTTGCTGAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGCCCTCAGCACC
TCCTCCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCT
TTCTGATCAGGAGGCTTCTTATGAATTAAACTGCCACCACCCCTCA

FIGURE 50

MERVTLALLLAGLTALLEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCKYKS
SQKQHSPVPEKAIPPLITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG
AGGACAGGGAGTCGAAGGAGGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGCAAGG
AGGAGACCCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGCTGGCAGAGATGAAGTCCAG
GGGCCCTGGCCTGCCTCTGCTGGCCCTGCTGGCAGTGGGAGGCTGGCCCTGCAGAG
CGGAGAGGAAAGCACTGGGACAAAATTGGGAGGCCCTGGACATGGCCTGGAGACGCCCTGA
GCGAAGGGGTGGGAAAGGCCATTGGCAAAGAGGCCGGAGGGCAGCTGGCTCTAAAGTCAGTGAG
GCCCTGGCCAAGGGACCAGAGAACAGCAGTGGCACTGGAGTCAGGCAGGTTCCAGGCTTGGCGC
AGCAGATGCTTGGCAACAGGTGGGAAGCAGCCATGCTCTGGAAACACTGGCACGGAGA
TTGGCAGACAGGCAAGAGATGTCATTGACACGGAGCAGATGCTGTCCGGCTCTGGCAGGGG
GTGCCTGGCACAGTGGTCTGGAAACTCTGGAGGCCATGGCATCTTGGCTCTCAAGGTGG
CCTTGGAGGCCAGGGCAAGGGCAATCCTGGAGGTCTGGGACTCCGTGGTCCACGGATAACCCG
GAAACTCAGCAGGCCAGCTTGAATGAATCCTCAGGGAGCTCCCTGGGTCAAGGAGGCAATGGA
GGGCCACCAAAACTTGGGACCAACACTCAGGGAGCTGTGGCCAGCCTGGCTATGGTCAGTGAG
AGCCAGCAACCAGAATGAAGGGTGCACGAATCCCCACCATCTGGCTCAGGTGGAGGCTCAGCA
ACTCTGGGGAGGCAGCGCTCACAGTCGGCAGCAGTGGCAGTGGCAGCAATGGTACAACAAAC
AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCGCAGCAG
TGGCGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGCAGTGAGT
CCTCCTGGGATCCAGCACCGCTCCTCCGGCAACCACGGTGGAGCGGGAGGAAATGGGATTCAAGG
CATAAACCCGGGTGTAAAAGCCAGGGAAATGAAGCCGGAGCAGGGAAATCTGGGATTCAAGG
CTTCAGAGGACAGGGAGTTCCAGCAACATGAGGGAAATAAGCAAAGAGGGCAATGCCTCTTG
GAGGCTCTGGAGACAATTATCGGGGCAAGGGTCAAGGGTCAAGGGTCAAGGGTCAAGGGTCAAGGG
GGTGGAGTCATACTGTGAACCTGAGACGTCTCTGGATGTTAACCTTGACACTTCTGGAA
GAATTAAATCCAAGCTGGTTTCACTCAACTGGGATGCCATAAACAAAGGACCAAGAGCTC
GCATCCCGTGACCTCCAGACAAGGAGCCACCAGATTGGATGGAGGCCACACTCCCTCTAA
AACACCACCCCTCATCACTAATCTCAGCCCTGGCCCTGAAATAACCTTAGCTGCCCAACAAA
AA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 52

MKFQGPLACLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLDALSEGVGKAIGKEAGGAAGSKVS
EALGQQGTREAVGTGVRQVPGFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGP
GHSGAWETSGGHGIFGSQGGLGGQQGQGNPGGLGTPWVHYPGNSAGSFGMNPQGAPWGQGGNGGPPNF
GTNTQGAVAQPGYGVASRNQNNEGCTNPPPSGSGGSSNSGGSGSQGSSGSGSNGDNNNGSSSGS
SSGSSSSGSSGGSSGGSSGGSSGNSGGSRGDSGSESSWGSSSTGSSSGNHGGGGNGHKGCEKPGNE
ARGSGESGIQGFRGQGVSSNMREISKEGNRLGGSGDNYRGQGSSWGSGGDAVGGVNTVNSETSPGM
FNFDTFWKNFKSKLGFINWDAINKDQRSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,
90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,
159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,
236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,
253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,
283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,
298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,
325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAAGAGGTTGTGTGGGACAAGCTGCTCCCACAGAAGGATGTCCTGCTGAGCCTGCCCTGG
CTGGGCCTCAGACCGGTGGCAATGTCCCCATGGCTACTCCTGCTGCTGGTTGTGGCTCCTGGCT
ACTCGCCCGCATCCTGGCTGGACCTATGCCCTATAACAACTGCCGCCGGCTCCAGTGTTC
CACAGCCCCAAAACGGAACTGTTTGGGTCACCTGGGCTGATCACTCCTACAGAGGAGGGC
TTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGTTACGGTATGGCTGGGTCCCAT
CATCCCCCTCATCGTTATGCCACCCCTGACACCATCCGGTCTATACCAATGCCCTAGCTGCCA
TTGCACCCAAGGATAATCTCTCATCAGGTCCTGAAGCCCTGGCTGGAGAAGGGATACTGGCTG
AGTGGCGGTGACAAGTGGAGCCACCGTCGGATGCTGACGCCCTCCATTCAACATCCT
GAAGTCCTATATAACGATCTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCAGCACCTGG
CCTCAGAGGGCAGCAGTCGTCGGACATGTTGAGCACATCAGCCTCATGACCTTGACAGTCTA
CAGAAATGCATCTCAGCTTGACAGCATTGTCAGGAGAGGCCAGTGAATATATTGCCACCAT
CTTGGAGCTCAGTGCCTTGTAGAGAAAAGAACGCCAGCATATCCTCCAGCACATGGACTTCTGT
ATTACCTCTCCATGACGGGCGGCGCTTCCACAGGGCTGCCGCTGGTGCATGACTCACAGAC
GCTGTCATCCGGGAGCGCGTCGCACCCTCCCACTCAGGTATTGATGATTTCAAAGACAA
AGCCAAGTCCAAGACTTGGATTTCATTGATGTCCTGCTGAGCAAGGATGAAGATGGAAAGG
CATTGTCAGATGAGGATATAAGAGCAGAGGCTGACACCTCATGTTGGAGGCCATGACACCACG
GCCAGTGGCCTCCTGGTCTGTACAACCTTGCAGGCACCCAGAATACCAGGAGCGCTGCCG
ACAGGAGGTGCAAGAGCTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGACGACCTGGCCC
AGCTGCCCTTCCGTACCATGTCGTGAAGGAGAGCCTGAGGTTACATCCCCAGCTCCCTCATC
TCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGCCGAGTCATCCCCAAGGCATTACCTG
CCTCATCGATATTATAGGGTCCATACAACCAACTGTGTGGCCGGATCCTGAGGTCTACGACC
CCTTCCGCTTGACCCAGAGAACAGCAAGGGAGGTCACCCTGCTTTATTCCCTTCTCCGCA
GGGCCAGGAACTGCATGGGCAAGGCGTTCGCCATGGGGAGATGAAAGTGGCTCTGGCGTGAT
GCTGCTGCACTTCCGGTCTGCCAGACCACACTGAGCCCCGAGGAAGCTGGAATTGATCATGC
GCGCCGAGGGCGGGCTTGGCTGCGGGTGGAGCCCTGAATGTAGGCTTGCAGTACTTCTGAC
CCATCCACCTGTTTTGCAAGATTGTCATGAATAACGGTGCTGTCAAA

FIGURE 54

MSLLSLPWLGRLPVAMSPWLLLLLVVGSWLLARI LAWTYAFYNNCRLQCFPQPPKRNWFWGHLG
LITPTEEGLKDSTQMSATYSQGFTVWLGPPIIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP
WLGEGLLSGGDKWSRHRMLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI
SLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRRFHAC
RLVHDFTDAVIRERRTLPTQGIDFFKDKAKSKTLDFIDVLLLSKDEDGKALSDEDIRAEADTF
MFGGHDTTASGLSWVLYNLRHPEYQERCRCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPFRRFPENSKGRSP
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN
VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCCATGGGACCAGTGAAACAGCTGAAGCGAATGTTGA
GCCTACTCGTTGATTGCAACTATCATGGTGTGTTGCACTTACCCGTGTTGCCT
TTTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTGCAGTCTTGGCATTGACGTGG
TACAGCCTTCCTTCATACCATTGCAAGGGATGCTGTGAAGAAGTGTGTTGCCGTGTCCTGC
ATAATTCATGCCAGTTTATGAAGCTTGGAAAGGCACTATGGACAGAAGCTGGTGGACAGTTT
GTAACATCTCGAACACCTCTGTCTTACAGACATGTGCCATTATCTTGCAGCAATGTGTTGCTT
GTGATTGAAACATTGAGGGTTACTTTGGAAGCAACAATACATTCTCGAACCTGAATGTCAGTA
GCACAGGGATGAGAAGTGGGTTCTGTATCTTGTGGAGTGGAAATCTCCTCATGTACCTGTTCTC
TCTGGATGTTGTCCTGAATTCCCAGTAAATACAAACCTATTCAAGCAACAGCAAAAAAAAAAAAA
AAA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCASFWWHNKGLALIFCILQSLALTWYSLSFIPFAR
DAVKKCFAVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

© 2002 University of Michigan Health System

FIGURE 57

CGGCTGAGCTGAGCCGAATGGCTCGAGGGCAGTGGAGCACCCAGCAGGCCAACATGCTCTGTCTGTGCCTG
TACGTGCGGTGATGGGAAGGCCAGACCGAGTTCACTTGTAGTCGAAGGGCTCTGGCAGCTGAAGTCC
ATTTCAGCTCAGTGCTTCATCCCAGGAATTCTCCACCTACCGCAGTGGAAAGCAGAAAATTGTACAAGCT
GGAGATAAGGACCTTGATGGCAGCTAGACTTTGAAGATTTCAGTCCAAAGATCATGAGAAAGCTGAGG
CTGGTGTAAAGATTTGGACAAAAAAGATGATGGCAGATTGACGCCAGGAGATCATGAGTCCCTGCGGGACTTG
GGAGTCAAGATACTGAAACAGCAGGCCAGAAAATTCTCAAGACATGGATAAAACGGCAGATGACCATGACTGG
AACGAGTGGAGAGACTACCACCTCTCCACCCGGTGGAAAACATCCCCAGATCATCCTACTTGAAGCATTCCAG
ATCTTGATGTGGTGAGAATCTAACGGTCCCGATGAGTTCACTGGAGGAGGGCAGACGGGATGTGGTGGAGA
CACCTGGCAGGAGGTGGGCAAGGGCGTATCCAGAACCTGACGCCAGGGCTGACAGGCTAAAGGTGCTCATG
CAGTCCATGCCCTCCAGCAACAACTGGCAGTGGCTTCACTCAGATGATTCGAGAAGGGAGGGCAGG
TCACTCTGGGGGCAATGGCATAACGCTCTCAAATTGCCCCGAAATCAGCCATCAAATTGCTATGAGCAG
ATCAAGGCCCTTGTGGTAGTGAACAGGAGACTCTGAGGATTACGGAGAGGCTTGCGAGGGTCTTGCAGGGCC
ATCGCCAGAGCAGCATCTACCCAATGGAGTCTCTGAAGACCCGGATGGCCTGCGGAAGACAGGCCAGTACTCAGGA
ATGCTGGACTGCCAGGAGGATCTGGCAGAGAGGGGTGGCCCTTCTACAAAGGCTATGCCCCAACATGCTG
GGCATCATCCCCATGGCCATGCCCTGAGCTACGAGACGCTCAAGAATGCTGGCTGAGCAGTATGAGT
AACAGCCGGACCCGGCTGTTGCTCTGGCCTGTGGCACCTGAGCTACGAGACGCTCAAGAATGCTGGCAGCTG
CCCCCTGCCCTAGTCAGGACCCGGATGCAGGGCAAGCCTTATTGAGGGCCTCCGGAGGTGACCATGAGCAGCCTC
TTCAAACATATCTGCCGACGGGGGGCTTCGGCTGTACAGGGGCTGGCCCAACTTCATGAAGGTATCCCC
GCTGTGAGCATCAGTACCTGGTCACTGAGAACCTGAGATCACCTGGCTGCACTCGCGTGAAGGGGAGGGC
CGCCCGCAGTGGACTCGTGCATCTGGCAGGCTGGGTGTGCAAGGATCTTGTGAAATGCTGCCAACACT
AAGCTGTCGAGCCAAAGCTGAAACCCCTGAGCCACCCGAGGGAGGTGGAGAGCTGGCAGGCCAGGGCTT
GTCTGCTGACCCAGACCCCTCTGTTGCTCAGGAAAGCAGGCTTCTGAGGAGGAGGAGGAGGAGGAGGAGG
CTCCGGCTCACATGTGAAGGACAGGACATTTCAGTGCCTGCAATAGTGAGCTGGAGGCTGGAGGCCGGCT
TAGTTCTCCATTTCACCCCTGCAAGGAGCTGTTGCCACGGCCCTGCCCTGGCTGCGCTGCATCTCCCTGTC
CCTCTGCTGCTGCCCTGCTGCTGAGGTAAAGGTGGGAGGGAGGCTACAGCCCACATCCCACCCCTCGCCAATCCC
ATAATCCATGATGAAAGGTGAGGTCAACGTGGCTCCAGGGCTGACTTCACCTGAGCCATGAGCAGCAACTTGG
TGTGAACGAAAGGAAAGGATCTGGCTGTGCTCACTGGCATCTGAGCCCTGCTGATGGCTGGGCTCTGGCATG
CTTGGGAGTGCAGGGGCTGGCTGGCTGGCTGCAAGAGGCAAGTGTGCTGGGCTCATGGTCTGAGCT
GGCCTGGACCCCTGTCAGGATGGGCCCCACCTCAGAACCAAACACTGTCCTGGCATGAGGGCAGTGGACCA
CCATGTTGAGGGCAAGGGCAGAGCTTGTGTTCTGGGAGGGAGGAAAGGATGTGGAGGCTTAATTATGG
ACTGTTGGGAAAGGGTTTGTCCAGAAGGAAAGGCCAACATGAGGACTTCTGCTTCCAGAGGAAGACGAG
GAGCAGGAGCTGGCTGACTGCTCAGAGTCTGTTCTGAGCCCTGGGGCTCTGCTCAACCCAGCAGGGCCAGC
GGGACCAAGCCACATCCACTTGTCACTGCTGGAACCTATTATTGGTATTATTGAACAGAGTTATGCT
AACTATTTTATAGATTGTTAAATAGCTGTCATTTCAAGTCTTTTATTATGTTATGTTATGTTATGTT
GATTGTAACCTCCCAAGGCCGCCCCAGTGGGATGGGAGGGAGGAGGAGAAGGGGGCTGGCTGAGTCACATCT
CTCCAGAGAAATTCTTGGGACTGGCAGAACAGGCCAGAACAGGCCAGAGGCAACGGCCCTGGCTCTTCTGGCAG
GTGGGGAGGGCTGGCCCTAGGATTCTGGGTTCTGACTGGGGCTGGGGAGAGAGAGGGAGGAGCT
AACCTTGAGGTGAACTCCAGTTATTCCTGCGCTGCGAGGGTTCTTATTCACCTTTCTGAATGTCAGGCAG
TGAGGTGCTCTCACTGTGAATTGTGTTGGGGGGCTGGAGGAGGGTGGGGGCTGGCTCGTCCCTCCAGC
CTTCTGCTGCCCTGCTTAACATGCCGGCAACTGGCACCTCACGGTTGCACCTCCATTCCACAGAACATGACCTGA
TGAGGAATCTCAATAGATGCAAAAGTCAATGCAAAATTGTTATATGACACATATAACTGGAGTCGTC
CAAATTAAGAAAGATGGACGTTAGAAGTTGTCATTAAAGCAGCCTCTAATAAAAGTTGTTCAAAGCTGAAAAA
AAA

FIGURE 58

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFTSTYRQWKQKIVQAGDKDLDG
QLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSILRDLGVKISEQQAEKILKSMDKNG
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMWWRHLVAGGG
AGAVSRTCTAPLDRLKVLQMVKHASRSNNMGIVGGFTQMIREGGARSILWRGNGINVLKIAPESAIIK
FMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIAQSSIYPMEVLKTRMALKTGQYSGMLDCARR
ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNAWLQHYAVNSADPGVFVLLACGTMSSTC
GQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV
VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATA CGCTGGAACCTTCCCCAGCCATGGC
TTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATTGGCTTGGTATTCAGGGAGACACTCCATCACAGTCACTACTGTCGCCTCAGCT
GGGAACATTGGGAGGATGGAATCCTGAGCTGCACHTTGAAACCTGACATCAAACCTTCTGATAT
CGTGATACAATGGCTGAAGGAAGGTGTTAGGCTGGTCCATGAGTTCAAAGAAGGCAAAGATG
AGCTGTCGGAGCAGGATGAAATGTTCAGAGGCCGGACAGCAGTGTGCTGATCAAGTGATAGTT
GGCAATGCCTCTTGC GGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
CATCACTTCTAAAGGCAAGGGAAATGCTAACCTTGAGTATAAAACTGGAGCCTCAGCATGCCGG
AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGC GGTGTGAGGCTCCCCGATGGTCCCC
CAGCCCACAGTGGCTGGCATCCAAAGTTGACCAGGGAGCCAACCTCTCGGAAGTCTCCAATAC
CAGCTTGAGCTGAAC TCTGAGAATGTGACCATGAAGGTTGTGCTGTGCTACAATGTTACGA
TCAACAAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGATATCAAAGTG
ACAGAAATCGGAGATCAAAGGGAGTCACCTACAGCTGCTAAACTCAAAGGCTCTGTGTTG
CTCTCTTCTTGCCATCAGCTGGCACTCTGCCTCTCAGCCCTTACCTGATGCTAAATAAT
GTGCCTTGGCCACAAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTCA
CACCAGATATGACCTAGTTTATATTCTGGGAGGAATGATCATATCTAGAAAGTCTGGAGTG
AGCAAACAAGAGCAAGAAACAAAAGAGCCAAAGCAGAAGGCTCAATATGAACAAGATAAAT
CTATCTCAAAGACATATTAGAAGTTGGAAAATAATTCA TGACTAGACAAGTGTGTTAAGA
GTGATAAGTAAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCGTGCTGT
CACCTGGGGAGTGAGAGGACAGGGATAGTGCATGTTCTGTCTGAATTTTAGTTATATGTG
TGTAATGTTGCTCTGAGGAAGCCCCTGGAAAGTCTATCCAAACATATCCACATCTTATATTCCAC
AAATTAAGCTGTAGTATGTACCTAAGACGCTGCTAATTGACTGCCACTCGCAACTCAGGGCG
GCTGCATTTAGTAATGGGTCAAATGATTCACTTTTATGATGCTTCAAAGGTGCCTGGCTTC
TCTTCCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTAGCATAAACAGAGCAGT
CGGGGACACCAGTTATAAATAAACTGAGCACCTTCTTTAAACAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAA

FIGURE 6o

MASLGQILFWSIISIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTPEPDIKLS
DIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTDAGTYKC
YIITSKGKGNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFPQPTVVVASQVDQGANFSEVS
NTSFELNSENVTMKVSVLYNVTINNTYSCKMIENDIAKATGDIKVTSEIIRRSHLQLLNASKSL
CVSSFFAISWALLPLSPYLMNK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTAGAACATCACCATGGCCAGCTATCCTTACCGGCAGGGCTGCCAGGAGCTGCAGGACAAG
CACCAAGGAGCCCCCTCCGGGTAGCTACTACCCCTGGACCCCCAATAGTGGAGGGCAGTATGGTAGT
GGGCTACCCCCCTGGTGGTGGTTATGGGGGCTCTGCCCTGGAGGGCTTATGGACCACCAGCTGG
TGGAGGGCCATGGACACCCAACTCCTGGGATGTTCCCCTCTGAAGCTCCAGGAGGACCATATG
GCGGTGCAGCTCCGGGGGCCCCCTATGGTCAGCCACCTCCAAGTTCCTACGGTGCCCAGCAGCCT
GGGCTTATGGACAGGGTGGCCCTCCCAATGTGGATCCTGAGGCCTACTCCTGGTTCCAGTC
GGTGGACTCAGATCACAGTGGCTATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
ATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTGACAAGACCAAGTCA
GGCCGCATCGATGTCACGGCTCTCAGCCCTGTGGAAATTCCAGCAGTGGAAAGAACCTCTT
CCAGCAGTATGACCGGGACCGCTCGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCC
AAATGGGCTACAACCTGAGCCCCAGTTCACCCAGCTCTGGTCTCCGCTACTGCCAACGCT
GCCAATCCTGCCATGCAGCTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA
GGCCTCCGGGAGAAGGACACAGCTGTACAAGGCAACATCCGCTCAGCTCGAGGACTTCGTCA
CCATGACAGCTCTCGGATGCTATGACCCAAACCATCTGTGGAGAGTGGAGTGCACCAGGGACCTT
TCCTGGCTCTAGAGTGAGAGAATGTGGACATCTTCTTCCCTGTCCCTCTAGAAGAAC
ATTCTCCCTTGCTTGATGCAACACTGTTCCAAAAGAGGGTGGAGAGTCCTGCATCATGCCACCA
AATAGTGAGGACGGGGCTGAGGCCACACAGATAGGGCCTGATGGAGGAGAGGATAGAAGTTGA
ATGTCTGATGGCCATGAGCAGTTGAGTGGCACAGCCTGGCACCAGGAGCAGGTCTTGTAAATGG
AGTTAGTGTCCAGTCAGCTGAGCTCCACCCCTGATGCCAGTGGTAGTGTTCATGGCCTGTACC
GTTAGTACCTGTGTTCCCTCACCAAGGCCATCCTGTCACAGGCCATTTCCTCAAAGTGGAAAT
CTGACCAAGCATGAGAGAGATCTGTCTATGGACCAGTGGCTGGATTCTGCCACACCCATAAAAT
CCTTGTGTGTTAACTCTAGCTGCCTGGGCTGCCCTGCTCAGACAAATCTGCTCCCTGGCCT
CTTGCCAGGCTCTGCCCTGCAGCTGGGACCCCTCACTTGCCCTGCCATGCTCTGCTCGGCT
TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTAATTGCATTTTTTC
ATTTGGGGCCAAAAGTCCAGTGAAATTGTAAGCTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSTGPGPYGGAAPGGPYGQPPPSSYGAQQPGLYQGGAPPNVDPEAYSWFQSVDSDH
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTSGRIDVYGFSAWKFIQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQLTEAFREK
DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-
66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCCGCGTGGCAGGGAGCTGCCTCCTGGGCCTGCTCCTGGCTGTCTTCATC
TCCCAGGCCTCTTGCCCGAGCATCGGTGTTGAGGAGAAAGTTCCAAAATTCGGGACC
AACTTGCCCTAGCTCGAACCTCCTCCACTGGCCCTCTAACCTGAACATCCGAGCCGC
TCTGGACCCCTAGGTCTAATGACTTGGCAAGGGTTCTGAAGCTCAGCGTGCCTCATCAGATG
GCTTCCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGCCCTCATCGTGGGGCTGCCTGCCATG
GATT CCTGGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGC GGCTGAGGACCGCCTGGGGGA
AGCGCTGCCTGAAGAACTCTCTTACCTCTCCAGTGCTGCGGCCCTCGCTCCGGCAGTGGCCCTT
TGCCTGGGGAGTCTTCTCCGATGCCACAGGCCTCTCACCTGAGGCTCACTCCTCCACCAGGAC
TCGGAGTCCAGACGACTGCCCGTTCTAACCTACTGGAGCCGGGGAAAAATCCTTCCAAACG
CCCTCCCTGGTCTCATCCACAGGGTTCTGCCTGATCACCCCTGGGTACCCGAATCCCAGTG
TGT CCTGGGGAGGTGGAGGCCCTGGGACTGGTGGGAACGAGGCCATGCCACACCTGAGGGA
ATCTGGGTATCAATAATCAACCCCCAGGTACCACTGGGAAATATTAAATCGGTATCCAGGAGG
CAGCTGGGGAAATATTAAATCGGTATCCAGGAGCAGCTGGGAAATATTAAATCGGTATCCAGGAG
GCAGCTGGGGAAATATTCACTATACCCAGGTATCAATAACCCATTCCCTGGAGTTCTCCGC
CCT CCTGGCTCTTCTGGAACATCCCAGCTGGCTTCCCTAACCTCCAAGCCCTAGGTGCAGTG
GGGCTAGAGCACGATAGAGGGAAACCCAACTTGGGAGTTAGAGTCTGCTCCGCCCTGCTG
TGTGGGCTCAATCCAGGCCCTGTTAACATGTTCCAGCACTATCCCCACTTTCA GTGCCTCCCC
TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AA

FIGURE 64

MQGRVAGSCAPLGLLLVCLHLPLGFARSIGVVEEKVSQNFGTNLPQLGQPSSTGPSNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPPAGGSAVQRWPPSWGLPAMDSWPPEDPWQMMAAAEDRLGEA
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLPRSNSILGAGGKILSQRP
PWSLIHRVLPDHPWGTLNPSVSWGGGPCTGWGTRPMPHPEGIWGINNQPPGTWGNINRYPGGS
WGNINRYPGGSGNINRYPGGSGNIHLYPGINNPFPFGVLRPPGSSWNIPAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCATCCCAGGAGCCAGTGGCCACTATGGGTC
TGGGCTGCCCTTGTCTCCTCTTGACCCCTTGGCAGCTCACATGGAACAGGGCCGGGTATGA
CTTGCAACTGAAGCTGAAGGAGTCTTCTGACAAATTCCCTCATGAGTCCAGCTCCTGGAA
TTGCTTGAAAAGCTCTGCCTCCCTCCATCTCCCTCAGGGACCAGCGTCACCCCTCCACCATGC
AAGATCTCAACACCATGTTGTCGAAACACATTGACAGCCATTGAAGCCTGTGTCCTTCTTGGCCC
GGGCTTTGGGCCGGGATGCAGGAGGCAGGCCCGACCCCTGTCTTCAGCAGGCCACCCCTC
CTGAGTGGCAATAATAAAATTGGTATGCTG

FIGURE 66

MGSGLPLVLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSYESSFLELLEKLCLLHLPSGTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCGGC
CAGGTGCCCGTCGCAGGTGCCCTGGCCGGAGATGCGTAGGAGGGCGAGCGCGAGAACCCCC
TTCCTCGGCCTGCCAACCCGCCACCCAGCCCATGGCGAACCCGGCTGGGCTGCTTCTGGCG
CTGGGCCTGCCGTTCTGCTGCCCGTGGGCCGAGCCTGGGGCAAATACAGACCAACTCTGC
AAATGAGAATAGCACTGTTGCCCTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGCTTCTCCCTCTGGCTGCCTGCTGGCTGTGGGCTG
GCACTGTTGGTGCGGAAGCTCAGGGAGAACCGGAGACGGAGGGCACCTACCGGCCAGTAGCGA
GGAGCAGTTCTCCATGCAGCCGAGGCCGGCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
GCCTGCCATCTAGTCCCTCTCCTGCATCTGTCTCCCTCATTGCTGTGACCTTGGGAAA
GGCAGTGCCCTCTGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAAGAAGGTACTTCAA
AGACTCTGCCCTGAGGTCAAGAGAGGATGGGCTATTCACTTTATATATTATATAAAATTAG
TAGTGAGATGTAaaaaaaaaaaaaaaa

FIGURE 68

MANPGLGLLLALGLPFLLARWGR~~A~~WGQIQTTSANENSTVLPSS~~T~~SSSDGNLRPEAITAIIVVFS
LLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQF~~S~~HAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

FIGURE 70

MGLFRGFVFLVLCLLHQSNSTFIKLNNNGFEDIVIVIDPSVPEDEKIIEQIEDMVTTASTYLFE
ATEKRFKKNVSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTECGEKG
EYIHFTPDLGGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRC
SAGISGRNRYVKCQGGSCLSRACRIDSTKLYGKDCQFFPDKVQTEKASIMFMQSIDS
VVEFCNEKTHNQEAPSLQNIKCNFRSTWEVISNSEDFKNTIPMVT
PPPPPVSLLKISQRIVCLVLDKSGSMGGKDRLNR
MNQAAKHFLQLQTENGWSWGMVHFSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSICSGIK
YAFQVIGELHSQDGSEVLLLTGDGEDNTASSCIDEVKQSGAIVHFA
LGRAADEAVIEMSKITGGSHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLES
KGTLNSNAWMNDTVIIDSTVGKDTFFLITWNSLPPSISLWDPSGTIMENFTVDATSK
MAYLSIPGTAKVGTWAYNLQAKANPETLTITVTSRAANSSVPPITVNA
KMNDVNSFPSPMIVYAEILQGYVPVLGANVTA
FIESQNGHTEVLELLDNGAGADSFKN
DGVSRYFTAYTENGRYSLKVR
HGGANTARLKLRP
LNRAAYIPGWVVNGEIEANPP
RPEIDEDTQTTLED
FSRTASGGAFV
VSQVPSLPLPDQYPPS
QITDLDATV
HEDIKIILTWT
TAPGDNFDVGKV
QRYII
RISASILD
LRDSFDDAL
QVNTTDLS
PKEANSKES
FAFKP
ENISEENATH
FIAIKS
IDKSNL
TSKVS
NIAQV
TLC
FIPQ
ANP
DDID
PTPT
PTPT
PDK
SHNS
GVN
ISTLV
LSV
IGSVVI
VN
FIL
STTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGGTGAAACCTGGGAGTAGAGTACTGACAGCAAAGACCGGAAAGACCATACTGCCGGCAGGGTGA
CAACAGGTGTCATCTTTGATCTCGTGTGGCTGCCTCTATTCAAGGAAGACGCCAAGGTAATTGACCCA
GAGGAGCAATGATGTAGCCACCTCTAACCTCCCTCTTGAAACCCCCAGTTATGCCAGGATTACTAGAGATGTC
ACTCAACCAAGCGGCCCTCGCCTAACCTTGTTGGAGAGAGAACCTTGTGGGCTGCCCTCTTAGCA
GTGCTCAGAAGTGAACCTGCCTGAGGGTGGACAGAAGAAAGGAAAGGTCCTCTTGCTGTTGGCTGCACATCAGGAA
GGCTGTGATGGAAATGAGGTGAAAACCTGGAGATTCACTTCAGTCATTGCTCTGCCCTGCAAGATCATCCTTTAA
AGTAGAGAAGCTGCTCTGTGTTGGCTTAACCTCAAGGGCAGAACCTGTTCTAGAAGGAATGGATGCAAGCAGCTC
CGGGGGCCCAAACGCATGCTCTGTGGCTAGCCCAGGGAAAGCCCTCCGTGGGGCCCGCTTGAGGGATGCC
ACCGGTTCTGGACGCTGGCTGAATGATGTGGCTGCCGGGCTGCCCTGCTGGGATTTCCGGGTGGT
CTGGCAGTGGCTCTCTGTGTCATCTGCTCTGTCATGACATGTTGGCTGCAACCCAAAAGGTGACGAGGAGCAG
GTTTGTGGCTCTCTGTGTCATCTGCTCTGTCATGACATGTTGGCTGCAACCCAAAAGGTGACGAGGAGCAG
CTGGCAGTGGCTCTGTGTCATCTGCTCTGTCATGACATGTTGGCTGCAACCCAAAAGGTGACGAGGAGCAG
CGCAACTACGTGAGCAGCTGAAAGGGCAGATGCAACAGCTCAAGGAGACTGCAGGAGAGGAGTGGAGCTCAGG
AAATGGGCACTACCAAGGCAGCGATGCTGCTGGCCTGGGCTGGACAGGAGCCCCCAGAGAAAACCCAGGCCGACCTC
CTGGCCTCCCTGCACTCGCAGGTGGACAAGGAGGTGAATGCTGGCAGTCAAGCTGCCAGAGTATGCAAGCAG
CCTTTGCACTAGCTTACTACAGAGGTGACCTAGCTGGACTGCCCTACCCGCCACCCGGAGGAGAACGCTGTG
AGGAAGGACAAGGGGATGAGTTGGTGAAGGCTTGAATCAGCTTGGAGACCTTGAACAACTCTGCAAGAGAACAGC
CCCAATCACCGCTTACAGGCCCTGATTATCATAGAAGGGATCTACCGAACAGAAAAGGGACAAGGGACATTGTAT
GAGCTCACCTCAAAGGGGACCAAAACACGAATTAAACGGCTCATCTTATTCGACCATTGACCCCATCATGAAA
GTGAAAAATGAAAAGCTCAACATGGCAACAGCTTACATGTTATGTCCTGCTGCAACAAAAGGGTGGACAAGTTC
CGGCAGTTCATGAGAATTTCAGGGAGATGTCATGAGCAGGATGGGAGAGTGGCTCATCTCAGTGTGTTACTTGGG
AAAGAAGAAATGAGTCAAAGGAAATCTGAAACACTTCAAGGCTCACAAGTCACTTACCTCAGGAACTTACCTCATC
CAGCTGAATGGAAATTCTCGGGAAAGGGACTTGTGTTGGACCCCTCTGGAAAGGGAGCAACGCTCTCTC
TTTTCTGTGATGTGGACATCTACTTCACATCTGAATTCTCAATACGTGTTAGGCTGAATACACAGCCAGGGAGAAG
GTATTTATCCAGTCTTTCAGTCAGTACAATCTGGCATAATATACGGCCACATGATGCACTCCCTCCCTGGAA
CAGCAGCTGGCTATAAAGAAGGAAACTGGATTGGAGACTTTGGATGGGAGTGGCTGTCAGTATCGTCAG
TTCATCATATAGTGGCTGATCTGGACATCAAAGCTGGGGAGAGGATGTCAGCTTGGACCTTATCGCAAGTATCTC
CACAGCAACCTCATGGTACGGACCCCTGCGAGGACTCTTCCACCTCTGGCATGAGAAGGCTGCATGGACGAG
CTGACCCCGAGCAGTACAAGATGTCATGCACTCAAGGCCATGACAGGAGCATCCACGGCCAGTGGCATGCTG
GTGTTCAAGCACGAGATAGAGGCTCACCTCGAAACAGAAACAGAAGACAACTAGCAACAAAAGCATGAACTCCAGA
GAAGGATTGGGGAGACACTTTCTCTCTGGCAATTACTGAAAGTGGCTGCAACAGAGAAAAGACTTCCATAAA
GGACGACAAGAAATTGGACTGATGGTCAAGAGTGGAGCTGAAAGGCTTGGATTTCTCTGTTGGCTTTTACACAGA
AATCAAAATCTCCCTTGCTGCAAGGAACTTGTGTTGCTTAACTGTCCTGGTTTCTGTTGTTGCTGTTG
AGATTATAAGCTTAATGGTGTGGAGGTTTGATGGTGTGTTACAATACACTGAGACCTGTTGTTGCTGTCATTGA
AATATTGATGTTAAAGAGCATTGTAAGGAAATCTGATGAGCAAGGCTATTTCTCTGTTGATGAAATGA
GCCTATCACAGGGCTGAGTCTGTTGCAATTACTGAAAGTGTGAAATGAGCAACAGGCTATGTTGACT
AGTGAAGTACATTAAGTAAAATGGACCAAGAAAAGAAAAGGAAACCTTCTGTTGATGAGTGGCTGTTGACT
TAACCAAAATAATCTGCTTATCTTTGGCTCTCTTAACTGTCCTGGTTTCTGTTGTTGTTGTTGTTGTTGACT
TTTTCTCTGTGAGTTAGTCGTTGCTTAACTTACACTTGTCAAGGCTTACAAGAGAGCACAAGTTGGCCTAC
ATTTTTATTTTTAAGAAGATACTTGGAGATGCAATTGAGAACTTCTGTTGCAAGGCTATGTTGACT
CCAAGGACATGCCAATGCTGATCTGTCAGGCAGTGAATGTCAGGCTATGAGACATAGGGAGGAATGGTTGACT
AATACAGACGCTACAGTACTTCTGTAAGAGTATTTCGAAGAGGAGCAACTGCAACACTGGAGGAAAAGAAAATGAC
ACTTTCTGTTTACAGAAAAGGAAACTCATTCACTGGTGAATCTGTCAGTGTACCTAAAAGTCAGAAACACATTT
CTCCTCAGAAGTAGGGACCGCTTCTACCTGTTAAATAACCAAGTATACCGTGTGAACCAAACAACTCTCTTTC
AAAACAGGGTGTCTCTGCTGGCTTCCATAAGAAGAAATGGAGAAAATATATATATATATATATATATTGT
GAAGAGTCAATCCATCTGCAAGAATCTAGTGGGAGTGGAGGTTTGCTACATGTTATCCACCCAGGCCAGGTGGAG
TAAGTAAATTCTGTTGAGCTCTGCTGTTGCTACAGTAAACTCATTGTTAAAAGCTTCAAGAAC
ATGCACTGAGCTAATTATCTCTTGAGTCTGCTGTTGCTACAGTAAACTCATTGTTAAAAGCTTCAAGAAC
ATTCAAGCTGTTGGTGTGTTAAAATGCAATTGATTTGACTGTTGAGTTGTTGATGAAATTAAATTAAAACACAGG
CCATGAATGGAAGGTGGTATTGACAGCTAATAAAATGATTGTTGAGGATGAA

FIGURE 72

MMMVRRGLLAWISRVVVLLVLLCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEWE
EQHRNYVSSLKRQIAQLKEELQERSEQRLRNGQQYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDK
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVRKDKRDELVEAIESAETLNNPA
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMVKNEKLNMAN
TLININVPLAKRVDKFQFMQNPREMCIEQDGDRVHLTVVYFGKEEINEVKGILENTSKAANFRNF
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGKKVFYPVLFSQY
NPGIIYGHDAVPPLQQQLVIKKETGFWRDFGFGMTQCQYRSDFINIGGFDLIDIKGWGGEDVHLYR
KYLHSNLIVVRTPVTRGLFHHLWHEKRCMDELTPEQYKCMCMQSKAMNEASHGQLGMLVFRHEIEAHL
RKQKQKTSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 73

GAGACTGCAGAGGGAGATAAAAGAGAGAGGGCAAAGAGGCAGCAAGAGATTGTCCTGGGATCCA
GAAACCCATGATAACCTACTGAAACACCGAATCCCCTGGAAGCCCACAGAGACAGAGACAGCAAGA
GAAGCAGAGATAAAATACACTCACGCCAGGAGCTCGCTCGCTCTCTCTCTCACTCCTC
CCTCCCTCTCTCTGCTGTCTAGCCTCTAGCCTCAAATTCCCAGTCCCTGCACCCCTC
CTGGGACACTATGTTGTTCTCCGCCCTCCTGCTGGAGGTATTTGGATCCTGGCTGCAGATGGGG
GTCAACACTGGACGTATGAGGGCCCACATGGTCAGGACCATTGCCAGCCTTACCTGAGTGT
GGAAACAATGCCAGTCGCCATCGATATTCAAGACAGACTGTGACATTGACCCTGATTGCC
TGCTCTGCAGCCCCACGGATATGACCAGCCTGGCACCGAGCCTTGGACCTGCACAACAATGCC
ACACAGTGCAACTCTCTGCCCTCACCCGTATCTGGGTGACTTCCCAGAAAATATGTAGCT
GCCAGCTCACCTGCACTGGGTCAAGAAAGGATCCCCAGGGGGTCAGAACACAGATCAACAG
TGAAGCCACATTGCAAGCTCCACATTGTACATTGACTCTGATTCTATGACAGCTTGAGTG
AGGCTGCTGAGAGGCCCTAGGGCCTGGCTGTCTGGCATCCTAATTGAGGTGGTGAGACTAAG
AATATAGCTTATGAACACATTCTGAGTCACCTGCATGAAGTCAGGCATAAAGATCAGAACACCTC
AGTGCCTCCCTCAACCTAAGAGAGCTGCTCCCCAACAGCTGGGCAGTACTCCGCTACAATG
GCTCGTCACAACCTCCCTGCTACCAAGAGCTGCTGGACAGTTTTATAGAAGGTCCAG
ATTTCAATGGAACAGCTGAAAAGCTTCAGGGACATTGTTCTCACAGAACAGGAGCCCTCAA
GCTTCTGGTACAGAACTACCGAGCCCTCAGCCTCTCAATCAGCGATGGCTTGCTTCA
TCCAAGCAGGATCCTCGTATACCACAGGTGAAATGCTGAGTAGGTGAGGAATCTGGTGGC
TGTCTGCCTCTCCTGGCTGTTATTCATTGCTAGAAAGATTGGAAGAACAGGCTGGAAAA
CCGAAAGAGTGTGGTCTCACCTCAGCACAAGCCACGACTGAGGCATAAATTCTCTCAGATA
CATGGATGTGGATGACTCCCTCATGCCATCAGGAAGCCTCTAAATGGGTGAGGATCTGG
CCAGAAACACTGTAGGAGTAGTAAGCAGATGTCCTCCCTGGACATCTCTAGAGAGGAAT
GGACCCAGGCTGTCATTCCAGGAAGAACTGCAGAGCCTCAGCCTCTCCAAACATGTAGGAGGA
ATGAGGAAATCGCTGTGTTAATGCAGAGANCAAACCTGTTAGTTGCAGGGGAAGTTGGG
ATATACCCCAAAGCTCTACCCCTCACTTATGCCCTTCCCTAGATATACTGCGGGATCT
CTCCTAGGATAAAGAGTTGCTGTTGAAGTTGATATTTGATCAATATATTGAAATTAAAG
TTCTGACTTT

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWPASYPECGNNAQSPIDIQTDSTFDPDLPALQ
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKGSPGGSEHQINSEAT
FAELHIVHYDSDSYDSLSEAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTPPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEEPSKLLV
QNYRALQPLNQRMVFASFIQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKIRKKRLENRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCCGCTGCCGCCGCTGCTGCTGCTCCTGGCGGCCTGGGACGGGAGTTCCCTGTGTC
TCTGGTGGTTGCCAACCTGCAAACATCACCTTCTATCCATCAACATGAAGAATGTCCCTACA
ATGGACTCCACCAGAGGGCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTCATCACAA
ATTGGCCCACCAGAGGGTGGCACTGACTACAGATGAGAAGTCCATTCTGTTGTCCTGACAGCTCC
AGAGAAAGTGGAAAGAAAATCCAGAAGACCTCCTGTTCCATGCAACAAATACTCCAATCTGA
AGTATAACGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGACCAACCAC
ACGCTGGTGCTCACCTGGCTGGAGGCCAACACTTTACTGCGTACACGTGGAGTCCTCGTCCC
AGGGCCCCCTGCCGTGCTCAGCCTCTGAGAAGCAGTGTGCCAGGACTTGAAAGATCAATCAT
CAGAGTTCAAGGCTAAATCATCTCTGGTATGTTGCCCATACTATTACCGTGTCTTT
TCTGTGATGGCTATTCCATCTACCGATAATCCACGTTGGCAAAGAGAAACACCCAGCAAATT
GATTTGATTATGAAATGAATTGACAAAAGATTCTTGTGCCCTGCTGAAAAAAATCGTGATTA
ACTTTATCACCTCAATATCTGGATGATTCTAAAATTCTCATCAGGATATGAGTTACTGGGA
AAAAGCAGTGTATCCAGCCTTAATGATCCTCAGCCCAGCGGAACCTGAGGCCCTCAGGA
GGAAGAGGAGGTGAAACATTAGGTATGCTCGCATTGATGGAAATTGGTGA
AAAACACGGAAGGTACTTCTCACCCAGCAAGAGTCCTCAGCAGAACAAATACCCCCGGATAAA
ACAGTCATTGAATATGAATATGATGTCAGAACCAACTGACATTGTCGGGGCCTGAAGAGCAGGA
GCTCAGTTGCAGGAGGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTGGCAG
TCTTGGGCCGAAACGTTACAGTACTCATACACCCCTCAGCTCCAAGACTTAGACCCCCCTGGCG
CAGGAGCACACAGACTCGGAGGGGGCCGGAGGAAGAGCCATCGACGACCCCTGGTCAGTGGGA
TCCCCAAACTGGCAGGCTGTGTATTCCCTCGCTGTCCAGCTTCGACCAGGATTCA
AGCCTCTGAGGGGATGGCTCGGAGAGGGTCTCTATCTAGACTCTATGAGGAGCCGGCT
CCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTCTATGGAGGAATGGGGTATA
TGTGAGATGGAAAACTGATGCAACACTCCTTGCCTTGTGAAACAAGTGAG
TCACCCCTTGATCCCAGCCATAAAGTACCTGGATGAAAGAAGTTTTCCAGTTGTCAGTGT
CTGTGAGAATTACTTATTCTTCTCTATTCTCATAGCAGTGTGATTGGTCTATGCA
GGTCTCTAACATGATGGTGGCCTCTGGAGTCCAGGGCTGGCCGGTTGTTCTATGCAGAGAA
AGCAGTCAATAATGTTGCCAGACTGGGTGCAGAATTATTCAAGGTGGGTGT

FIGURE 76

MSYNGLHQRFKELKLLTLCSISSQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIY
SNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRRAQPSEKQCARTLK
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILIGNEFDKRFVPAEK
IVINFITLNISDDSKISHQDMSSLGKSSDVSSLNDPQPSGNLRPPQEEEVKHTMLGYASHIMEIFC
DSEENTEGTSLTQQESLSRTIPPDKTVIEYEYDVRTTDICAGPEEQELSLOQEEVSTQGTLLESQA
ALAVLGPQLQYSYTPQLQDLDPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRLCIPSLSSFDQDS
EGCEPSEGDGGLGEEGLLSRLYEEPAPDRPPGENETYLMQFMEEWGLYVOMEN

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTCCCCAGGTCTGGCATCCTGCACTTGCTGCCCTCTGACAC
CTGGGAAGATGGCCGGCCCCTGGACCTTCACCCCTCTGTGGTTGCTGGCAGCCACCTGATC
CAAGCCACCCCTCAGTCCCCTGCAGTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAAGCTGAC
ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC
GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGAGCCTGGTGAACACCGTCTGAAGCACATC
ATCTGGCTGAAGGTCACTCACAGCTAACATCCTCCAGCTGCAGGTGAAGCCCTGGCCAATGACCA
GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA
TCGTGGAGTTCCACATGACGACTGAGGCCAAGCCACCATCCGCATGGACACCAGTCAAGTGGC
CCCACCCGCTGGCCTCAGTGAAGTGTGCTGCCACCAGCCATGGGAGCCTGCGCATCCAACGTGTGA
TAAGCTCTCCTTCTGGTGAACGCCCTAGCTAACGAGGTCACTGAAACCTCTAGTGCCTACCCCTGC
CCAATCTAGTAAAAACCAAGCTGTGCTCCGTATCGAGGCTTCCTCAATGGCATGTATGCAGAC
CTCCTGCAGCTGGTGAAGGTGCCCTTCCCTCAGCATTGACCGTCTGGAGTTGACCTTGTGA
TCCTGCCATCAAGGGTACACCATTCAAGCTCTACCTGGGGCCAAGTTGGACTCACAGGGAA
AGGTGACCAAGTGGTCAATAACTCTGCAGCTCCCTGACAATGCCACCCCTGGACAACATCCCG
TTCAGCCTCATCGTGAGTCAGGACGTGGTAAAGCTGCAGTGGCTGCTGTGCTCTCCAGAAGA
ATTCAATGGTCCCTGGTGGACTCTGTGCTTCCCTGAGAGTGCCTGCGCATGGTCAAGTCAGCATCGGGC
TGATCAATGAAAAGGCTGCAGATAAGCTGGATCTACCCAGATCGTAAGATCCTAACTCAGGAC
ACTCCCAGTTTATAGACCAAGGCCATGCCAAGGTGGCCAAGTGTGCTGGAGATGTT
TCCCTCCAGTGAAGCCCTCCGCCCTTGTCAACCTGGCATCGAACGCCAGCTCGGAAGCTCAGT
TTTACACCAAAGGTGACCAACTTAACTCAACTTGAATAACATCAGCTCTGATGGATCCAGCTG
ATGAACCTCTGGATTGGCTGGTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA
CTCCATCCTGCTGCCGAACCAGAAATGGCAAATTAAAGATCTGGGTCCCTGGCATGTCATTGGTGAAGG
CCTTGGGATTGAGGGCAGCTGAGTCCTCACTGACCAAGGATGCCCTTGCTTACTCCAGCCTCC
TTGTGGAAACCCAGCTCCTGTCTCCAGTGAAGACTGGATGGCAGCCATCAGGGAAAGGCTGG
GTCCCAGCTGGGAGTATGGGTGTGAGCTATAGACCATCCCTCTGCAATCAATAAACACTTG
CCTGTGAAAAA

2025 RELEASE UNDER E.O. 14176

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKDHNATSILQQPLLSAMREK
PAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE
FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDTIQLYLGA
KLDQSQGVTKWFNNSAASLTMP TLNDNIPFSLIVSQDVVKAAVAAVLSPEEFMVLLDSVLPESA
HRLKSSIGLIN EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSEALRPLFTLG
IEASSEAQFYTKGDQLILNLNNSDRIQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKL
RGVPVSLVKALGFEEAAESSLTKDALVLTPASLWKPSVPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCACTGAGAGGTCTGCCATGGCCTCT
CTTGGCCTCCAACTTGTGGCTACATCCTAGGCCTCTGGGCTTTGGCACACTGGTTGCCAT
GCTGCTCCCAGCTGGAAAACAAGTCTTATGTCGGTGCAGCATGTGACAGCAGTTGGCTTCT
CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATACCCAGTGTGACATCTATAGC
ACCCTCTGGCCTGCCGCTGACATCCAGGCTGCCAGGCCATGATGGTGACATCCAGTGAAT
CTCCTCCCTGGCCTGCATTATCTCTGTTGGCATGAGATGCACAGTCTTCTGCCAGGAATCCC
GAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTCATCCTTGGAGGCCTCTGGGATTC
ATTCCTGTGCCCTGGAATCTCATGGGATCCTACGGACTTCACTCACCACGGTGCCTGACAG
CATGAAATTGAGATTGGAGAGGCTTTACTTGGCATTATTCTCCCTGTTCTCCCTGATAG
CTGGAATCATCCTCTGCTTTCCCTGCTCATCCCAGAGAAATCGCTCCAACACTACGATGCTAC
CAAGCCAACCTCTGCCACAAGGAGCTCCAAGGCCTGGTAACCTCCAAAGTCAAGAGTGA
GTTCAATTCTACAGCCTGACAGGTATGTGTGAAGAACCAAGGGCCAGAGCTGGGGGTGGCTG
GGTCTGTAAAAACAGTGGACAGCACCCCGAGGGCCACAGGTGAGGGACACTACCAACTGGATCGT
GTCAGAAGGTGCTGAGGATAGACTGACTTGGCATTGGATTGAGCAAAGGCAGAAATGGG
GCTAGTGTAAACAGCATGCAGGTTGAATTGCCAAGGATGCTGCCATGCCAGCCTTCTGTTCC
TCACCTTGCTGCCCTGCCCTAAGCCCCAACCTCAACTGAAACCCATTCCCTAACCCA
GGACTCAGAGGATCCCTTGCCCTCTGGTTACCTGGACTCCATCCCCAACCCACTAACCA
TCCCACGTGACTGCCCTCTGTGATCAAAGACCCCTCTGGCTGAGGTTGGCTTAGCTCATT
GCTGGGATGGAGAGGAGAACAGTGGCTTGTGGCATTGCTCTAACCTACTTCTAACGCTTC
CCTCCAAAGAAACTGATTGCCCTGGAACCTCCATCCACTCTGTTATGACTCCACAGTGTCCA
GACTAATTGTGCATGAACTGAAATAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAAGGCAGCCTGGACATTAAAAAAATA

DNA sequence

FIGURE 8o

MASLGLQLVGYILGLLGLLGTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAQAMMVTTSSAISSLACIISVVMRCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPVAWNLHGIILRDFYSPLVPSMKFEIGEALYLGISSLFSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVKEFNSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCACACCGTCCGGCCCTCTCCCTCTGCTGGACCTTCCTCGTCTCCATCTCTCCCTCC
CCCGCGTTCTTTCCACCTTCTCTCCACCTAGACCTCCCTCCGTCCCTGCCCTCCT
GCCCACCGCTGCTCCCTGGCCCTCTCGACCCGCTCTAGCAGCAGACCTCCGGGGCTGTGG
GTTGATCTGTGGCCCTGTGCCCTCGTCTTCGTCCTCCCTCCGACTCCGCTCCCG
ACCAGCGGCCTGACCCGGAAAGGATGTTCCCGAGGTGAGGGTCCCTCCTCCCTGCTGGGA
CTCGCGCTGCTCTGGTTCCCCCTGGACTCCCACGCTCGAGCCGCCAGACATGTTCTGCC
CCATGGGAAGAGATACTCCCCCGCGAGAGCTGGCACCCCTACTTGGAGCCACAAGGCCTGATGT
ACTGCCCTGCGCTGTACCTGCTCAGAGGGGCCCATGTGAGTTGTTACCGCCTCCACTGTCCGC
GTCCACTGCCCTCAGCCTGTGACGGAGCACAGCAATGCTGTCCCAAGTGTGTGGAACCTCAC
TCCCTCTGGACTCCGGGCCACCAAAGTCTGCCAGCACACGGGACCATGTACCAACACGGAG
AGATCTCAGTGCCTCATGAGCTGTTCCCTCCCGCCTGCCAACCAAGCTGTGTCCCTGCAGCTGC
ACAGAGGGCCAGATCTACTGCCCTCACAAACCTGCCCGAACCAAGGCTGCCAGCACCCCTCC
ACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGCAAGTGAGCAATCGGATGAAGAGGACA
GTGTGCAGTCGCTCCATGGGTGAGACATCCTCAGGATCCATGTTCCAGTGATGCTGGAGAAAG
AGAGGCCGGCACCCAGCCCCACTGCCCTCAGGCCCTCTGAGCTTCATCCCTGCCACTT
CACAGACCAAGGGAGCAGGCAGCACAACTGTCAAGATCGCCTGAAGGAGAAACATAAGAAAGCCT
GTGTGCATGGCGGAAGACGTACTCCCACGGGAGGTGTGGCACCCGCCCTCCGTGCCCTCGC
CCCTGCCCTGCATCCTATGCCACCTGTGAGGATGCCGCCAGGACTGCCAGCGTGTGACCTGTCC
CACCGAGTACCCCTGCCGTACCCCGAGAAAGTGGCTGGGAAGTGCTGCAAGATTGCCAGAGG
ACAAAGCAGACCCGGCACAGTGAGATCAGTTCTACCAAGGTGTCCAAGGCACCGGGCCGGTC
CTCGTCCACACATCGGTATCCCAAGCCCAGACAACCTGCGTCGCTTGCCCTGAAACACGAGGC
CTCGGACTTGGTGAGATCTACCTCTGGAAGCTGGTAAAGATGAGGAAACTGAGGCTCAGAGAG
GTGAAGTACCTGGCCAAGGCCACACAGCCAGAAATCTCCACTTGACTCAGATCAAGAAAGTCAG
GAAGCAAGACTCCAGAAAGAGGCACAGCACCTCCGACTGCTCGCTGGCCCCACGAAGGTCACT
GGAACGTCTCCTAGCCCAGACCCCTGGAGCTGAAGGTACGCCAGTCCAGACAAAGTGACCAAG
ACATAACAAAGACTACAGTGCAGATATGAGCTGTATAATTGTGTTATTATATATTAATAAAA
TAAGAAGTTGCATTACCCCTCAAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPPVHCQPVTEPQQCCPKCVEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELP
PSRLPNQCVLCSCTEGQIYCGLTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPGTAPTGLSAPLSFI PRHFRPKGAGSTTVKIVLKEKKACVHGGKTYSHGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTEYPCRHPEKVAGKCCKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIYLWKLVDDETEAQRGEVPGPRPH
SQNLPLSDQESQEARNPERGTLPTARWPPRRSLERLPSPDPGAEGHGQSRQSDQDITKT

Signal peptide:
amino acids 1-25

2025 RELEASE UNDER E.O. 14176

FIGURE 83

GACAGCTGTGTCGATGGAGTAGACTCTCAGAACAGCGCAGTTGCCCTCGCTCACGCAGAGCCTCTCC
GTGGCTTCCGCACCTTGAGCATTAGGCAGTTCTCCTCTCTAATCCATCCGTACACCTCTCCTGTCA
TCCGTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTCTCATGCTCAGTTGGTCTGAGTC
TCCTCAAGCTGGATCAGGGCAGTGGCAGGTGTTGGCCAGACAAGCCTGTCAGGCCTTGGTGGGGAG
GACGCAGCATTCTCCTGTTCTGCTCCTAAAGACCAATGCAGAGGCCATGGAAGTGCCTTCTCAGGGG
CCAGTTCTAGCGTGGTCCACCTCTACAGGGACGGAGGCCATTATGCAGATGCCACAGTATC
AAGGCAGGACAAAATGGTGAAGGATTCTATTGCGGAGGGCGCATCTCTGAGGCTGGAAAACATTACT
GTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCCCAGTCTTAACAGAAGGCCATCTGGAGCT
ACAGGTGTCAGCACTGGGCTCAGTTCTCATTTCCATCACGGGATATGTTGATAGAGACATCCAGCTAC
TCTGTCAGTCCTCGGGCTGGTCCCCGGCCACAGCGAACAGTGGAAAGGTCCAAGGACAGGATTGTCC
ACAGACTCCAGGACAAACAGAGACATGCATGGCCTGTTGATGTTGAGATCTCTGACCGTCCAAGAGAA
CGCCGGGAGCATATCCTGTCATGCGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG
GAGATACCTTTTCGAGCCTATATCGTGGCACCTGGCTACCAAAGTACTGGGAATACTCTGCTGTGCC
TTTTTGGCATTGTTGACTGAAGATTTCTTCCAAATTCCAGTGGAAATCCAGGCGGAACGGACTG
GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCAACACGCAGTGGAGGTGACTCTGGATCCAG
AGACGGCTACCCGAAGCTCGCTTGTGATCTGAAAACGTAAACCCATAGAAAAGCTCCCCAGGAGGTG
CCTCAGTCTGAGAAGAGATTACAAGGAAGAGTGTGGTGGCTCTCAGAGTTCCAAGCAGGGAAACATTA
CTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGCGTGGAGTGTGCCGGATGATGTGGACAGGAGGA
AGGAGTACGTGACTTTGTCTCCGATCATGGTACTGGTCTCAGACTGAATGGAGAACATTGTATTTC
ACATTAATCCCCTTTATCAGCGTCTCCCCAGGACCCACCTACAAAATAGGGTCTCCTGACTA
TGAGTGTGGGACCATCTCTTCAACATAATGACCAGTCCCTATTATACCTGACATGTCGGTTG
AAGGCTTATTGAGGGCCTACATTGAGTATCGCTCTATAATGAGCAAAATGGAACCTAGTCATCTGC
CCAGTCACCCAGGAATCAGAGAAAGAGGCCCTTGGCAAAGGGCTCTGCAATCCAGAGACAAGAACAG
TGAGTCCTCCTCACAGGAACCACGCCCTCTCCCAAGGGTGAATTAGGATGAATCACATCCCACAT
TCTTCTTAGGGATATAAGGTCTCTCTCCAGATCCAAAGTCCGCAGCAGCCGCCAAGGTGGCTTCCA
GATGAAGGGGACTGGCTGTCCACATGGAGTCAGGTGTCAGGCTGCCCTGAGCTGGAGGGAAAGAAGG
CTGACATTACATTAGTTGCTCTCACTCCATCTGCTAAGTGATCTTGAATACACCTCTCAGGTGAAG
AACCGTCAGGAATTCCCATCTCACAGGCTGGTGTAGATTAAAGTAGACAAGGAATGTGAATAATGCTTAG
ATCTTATTGATGACAGAGTGTATCTAATGGTTGTTCAATTACACTTCAGTAAAAAAA

FIGURE 84

MALMLSIVSLLKLGGQWQVFGPDKPVQALVGEDAASFCLSPKTNAAEAMEVRFFRGQFSVVH
LYRDGKDQPFMOMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYYQKAIWELQ
VSALGSVP LISITGYVDRDIQLLCQSSGWPRPTAKWKGPGQGDLSTDRTNRDMHGLFDVEISL
TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
FQWKIQAEQDWRRKHGQAEQELRDARKHAVEVTLDPETAHPKLCVSDILKTTHRKAPEVPHSEKRF
TRKSVVVASQSFQAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTILSPDHGYWVRLNGEHLYFT
LNPRFISVFPPRTPTKIGVFLDYECGTISFFNINDQSLIYTLTCREFGLLRPYIEYPSYNEQNGT
PIVICPVTQESEKEASWQRASAIPETSNSESSSQATTPFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCCCAGACATGCTGCTGCTGCTGCTGCCCT
GCTCTGGGGAGGGAGAGGGCAGAAGGACAGACAAGTAAACTGCTGACGATGCAGAGTTCCGTGA
CGGTGCAGGAAGGCCTGTGTCCATGTGCCCTGCTCCTCTCCTACCCCTCGCATGGCTGGATT
TACCCGGCCCAAGTGTGCTACTGGTCCGGGAAGGGCCAATACAGACCAGGATGCTCC
AGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGAGGAGACTCGGGACCGATTCCACCTCCTG
GGGACCCACATACCAAGAATTGCACCCCTGAGCATCAGAGATGCCAGAAGAAGTGATGCCGGAGA
TACTTCTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACCGGCTCTGTGAA
TGTGACAGCCTTGACCCACAGGCCAACATCCTCATCCCAGGCACCCCTGGAGTCCGGCTGCC
AGAATCTGACCTGCTCTGTGCCCTGGGCTGTGAGCAGGGGACACCCCTATGATCTCCTGGATA
GGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCCGCTCCTCGTGCTCACCTCATCCCACA
GCCCGAGGACCATGGCACCGCCTCACCTGTCAGGTGACCTTCCCTGGGAGCGTGA
ACAAGACCGTCCATCTCAACGTGCTCACCGCCTCAGAACATTGACCATGACTGTTCAAGGA
GACGGCACAGTATCCACAGTCTGGAAATGGCTATCTGTCACTCCAGAGGGCCAGTCT
GCGCCTGGCTGTGCAGTTGATGCACTGAGCAATCCCGCTGCCAGGCTGAGCTGG
GAGGCTGACCTGTGCCCTCACAGCCCTCAAACCCGGGGTGCTGGAGCTGCC
CTGAGGGATGCAGCTGAATTCACCTGCAGAGCTCAGAACCCCTCTCGGCTCAGCAGGT
GAACGCTCCCTGCAGAGCAAAGCCACATCAGGAGTGA
CTCACAGGGGTGGTCGGGGAGCTGGAG
CCACAGCCCTGGTCTCCTGCGTCATCTCGTTGAGGTCTGCAGGAAGAAA
TCGGCAAGGCCAGCAGCGGGGTGGAGATACGGGCTAGAGGATGCAAACGCTGT
CAGGGTTCA
AGCCTCTAGGGGCCCTGACTGAACCTGGCAGAACAGACTCCCCAGACCAGCCT
CCCCAG
CTTCTGCCGCTCCTCAGTGGGGAGGAGGCTCCAGTATGCATCCCTCAGCT
CAGGCTGAGCTGGAG
AAGCCTGGACTCGCGGGACAGGAGGCCACTGACACCGAGTACTCG
GAGGATCAAGATCCACAG
ATGAGAAACTGCAGAGACTCACCCCTGATTGAGGGATCACAGCC
CTCCAGGCAAGGGAGAAGTCA
GAGGCTGATTCTGTAGAATTAAACAGCCCTAACGTGATGAGCTATGATA
ACACTATGAATTATG
TGCAGAGTGAAAGCACACAGGCTTAGAGTCAAAGTATCT
CAAACCTGAATCCACACTGTGCC
TCCCTTTATTTTTAACTAAAAGACAGACAAATTCTA

FIGURE 86

MLLILPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSYPSHGWIYPGPVVHGYWFREG
ANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLISRARRSDAGRYFFRMEKGSIKWNY
KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCSPWACEQGTPPMISWIGTSVSPLDPSTTRS
SVLTLIPQPQDHGTSLTCQVTFPGASVTTNKTVHLNVSYPPQNLMTVFQGDGTVSTVLNGSSL
SLPEGQSLRLVCAVDADSNPPARLSLSWRGLTLCPSQPSNPVGVLELPWVHLRDAAEFTCRAQNP
LGSQQVYLNVSLQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGDTGIE
DANAVRGSAQGPLTEPWAEDSPPDQPPPASARSSVGEHELQYASLSFQMVKPWDSRGQEATDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGTGAAGGAGCTCTGTAC
CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTCCTGCTGTTTC
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTCAAGGAATGGACCTGT
TCTTCGTCTCCATCTCTGCCAGAAGCTGAAAGGAAATCAAAGACGAATGTCTAGTGCATTG
TGGCCTGTATTTCTCCGCACTGAGAATGGTTATCTACCAGACCTCTGTGACATGACCTCTG
GGGGTGGCGGCTGGACCCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGAAAGTGCACGGT
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGGACGGCAACTGGC
CAACTACAACACCTTGATCTGCAGAGGCCACGAGCGATGACTACAAGAACCTGGCTACT
ACGACATCCAGGCCAAGGACCTGGCATCTGGCACGTGCCAATAAGTCCCCATGCAGCACTGG
AGAAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTCCAGACACTGGACATAATCT
GTTTGGCATCTACCAGAAATATCCAGTGAAATATGGAGAAGGAAAGTGTGGACTGACAACGGCC
CGGTGATCCCTG~~GGT~~CTATGATTTGGCAGCGCCAGAAAACAGCATCTTATTACTCACCTAT
GGCCAGCGGAATTCACTGCGGGATTGTTCAGTTCAGGGTATTTAATAACGAGAGAGCAGCCAA
CGCCTTGTGTGCTGGAATGAGGGTCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG
GATAACTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTTCTGGTTTGATGGAGTGGATAT
GGAACTCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
TTGAGAGTTTGTGGAGGGAACCCAGACCTCTCCCAACCATGAGATCCCAAGGATGGAGAA
CAACTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAATCATATTGACTCAAGA
AAAAAA

FIGURE 88

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRSSQQGSKADYPEGDGNWANYNTFGSAEAAT
SDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
EGKCWTDNGPVI PVVYDFGDAQKTASYYSPYQREFTAGFVQFRVFNNERAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPQQCGDFSGFDWSGYGTHVGYSSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTGTCGGCTTGCAGGGAGACTTCAGGAGTCGCTGTCTGAACCTCCAGCCTCAGAGAC
CGCCGCCCTTGTCGGCGAGGGCCATGGGCCGGTCTCAGGGCTGTGCCCTCGCTCCTGACG
CTCCTGGCGCATCTGGTGGTCGTACACCTTATTCTGGTCCCAGGACAGCAACATACAGGCCTG
CCTGCCTCTCACGTCACCCCCAGGGAGTATGACAAGCAGGACATTCAAGCTGGTGGCCGCGCTCT
CTGTCACCCTGGGCCTCTTGCACTGGAGCTGGCGGTTCTCAGGAGTCTCCATGTTCAAC
AGCACCCAGAGCCTCATCTCCATTGGGCTCACTGTAGTGCATCCGTGGCCCTGTCCCTTCAT
ATTTCGAGCGTTGGGAGTGCACACTACGTATTGGTACATTTGTCTTCTGCAGTGCCCTCCAGCTG
TCACTGAAATGGCTTATTGTCACCGTCTTGGGCTGAAAAAGAAACCTTCTGATTACCTCA
TGACGGGAACCTAAGGACGAAGCCTACAGGGCAAGGCCGTTCGTATTCTGGAAAGAAGGAAG
GCATAGGCTCGGTTTCCCTCGAAACTGCTCTGCTGGAGGATATGTGTTGAAATAATTACG
TCTTGAGTCTGGGATTATCCGATTGTATTAGTGTGTTGTAATAAAATATGTTTAGTAACA
TTAAGACTTATACAGTTAGGGACAATTAAAAAAAAAAA

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVVITLFWRSRDSNIQACIPLTFTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSALPAVTEMALFV
TVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

FIGURE 91

CTGGGACCCGAAAAGAGAAGGGAGAGCAGGGAGAGCAGGGAGGAAGGATGCAACTGAC
TCGCTGCTGTTCGTGTTCCTGGTGAGGGTAGCCTATCTGGTCATCTGTGCCAGGATGATG
GTCCTCCCGCTCAGAGGACCCCTGAGCGTGATGACCACGAGGGCAGCCCCGGGGCGCT
CGGAAGCGGGGCCACATCTCACCTAAGTCCC GCCCATGGCCAATTCCACTCTCCTAGGGCTGCT
GGCCCCGCTGGGGAGGCTTGGGCATTCTGGCAGCCCCCAACCGCCCGAACACAGCCCC
CACCCCTCAGCCAAGGTGAAGAAAATCTTGGCTGGGCACTTCAACTCCAAACATCAAGACGGTG
GCCCTGAACCTGCTCGTCACAGGGAAAGATTGTGGACCATGGCAATGGACCTTCAGCGTCCACTT
CCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCCTCGTCCCCCAGTAAAGCTGTAG
AGTTCCACCAAGGAACAGCAGATCTTCATCGAAGCCAAGGCCCTCCAAAATCTCAACTGCCGGATG
GAGTGGGAGAAGGTAGAACGGGCCGCCGACCTCGCTTGCACCCACGACCCAGCCAAGATCTG
CTCCCGAGACCACCGCTCAGAGCTCAGCCACCTGGAGCTGCTCCAGCCCTCAAAGTCGTCTGTG
TCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAATAC
CATAGTGAACCCCTACTACCCATCTGGGTGACCCGGGCAGGCCACAGAGGCCAGGCCAGGGC
TGGAGGACAGGCCCTGCCCATGCAGGAGACCATCTGGACACCAGGGCAGGGAAAGGGGTTGGCCTC
AGGCAGGGAGGGGGGTGGAGACGAGGAGATGCCAAGTGGGCCAGGGCAAGTCTCAAGTGGCAG
AGAAAGGGTCCAAGTGTGGTCCCAACCTGAAGCTGTGGAGTGAATAGATCACAGGAGCACTGG
AGGAGGAGTGGCTCTGTGCAGCCTCACAGGGCTTGCCACGGAGCCACAGAGAGATGCTGGG
TCCCCGAGGCCCTGTGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGAGGAAGCTAAC
CCTTGGTTCTGCCATCCTGAGGAAGATAGCAACAGGGAGGGGAGATTCACTCAGTGTGGACA
GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCCTAGGAGCCAGTCAGCAGGGTGGGG
GCCAGAGGAGCTCCAGCCCTGCCAGTGGGCCCTGAGCCCCCTGTGCTGAGCATGG
CATGAGGCTGAAGTGGCAACCTGGGTCTTGATGTCTGACAGATTGACCATCTGTCTCCAGC
CAGGCCACCCCTTCCAAAATCCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATTGCTGATG
GCACACCCATCCTTAAGCTAACAGACAGGAGCATTGTGGCCTCCACACTAACGGCCACAGCCCAC
CGCGTGTGTGTGTCCTCTCCACCCCAACCCCTGCTGGCTCTGGAGCATCCATGTCCCG
GAGAGGGGTCCTCAACAGTCAGCCTCACCTGTCAGACCCGGGTTCTCCGGATCTGGATGGCG
CGCCCTCTCAGCAGCGGGCACGGTGGGGGGGGCCAGAGCATGTGCTGGATCTGTT
TGTGTGTCTGTGTGGTGGGGAGGGAGGGAAAGTCTGTGAAACCGCTGATTGCTGACTTT
TGTGTGAAGAATCGTGTCTGGAGCAGGAATAAGCTTGGCCCGGGCA

900
899
898
897
896
895
894
893
892
891
890
889
888
887
886
885
884
883
882
881
880
879
878
877
876
875
874
873
872
871
870
869
868
867
866
865
864
863
862
861
860
859
858
857
856
855
854
853
852
851
850
849
848
847
846
845
844
843
842
841
840
839
838
837
836
835
834
833
832
831
830
829
828
827
826
825
824
823
822
821
820
819
818
817
816
815
814
813
812
811
810
809
808
807
806
805
804
803
802
801
800
799
798
797
796
795
794
793
792
791
790
789
788
787
786
785
784
783
782
781
780
779
778
777
776
775
774
773
772
771
770
769
768
767
766
765
764
763
762
761
760
759
758
757
756
755
754
753
752
751
750
749
748
747
746
745
744
743
742
741
740
739
738
737
736
735
734
733
732
731
730
729
728
727
726
725
724
723
722
721
720
719
718
717
716
715
714
713
712
711
710
709
708
707
706
705
704
703
702
701
700
699
698
697
696
695
694
693
692
691
690
689
688
687
686
685
684
683
682
681
680
679
678
677
676
675
674
673
672
671
670
669
668
667
666
665
664
663
662
661
660
659
658
657
656
655
654
653
652
651
650
649
648
647
646
645
644
643
642
641
640
639
638
637
636
635
634
633
632
631
630
629
628
627
626
625
624
623
622
621
620
619
618
617
616
615
614
613
612
611
610
609
608
607
606
605
604
603
602
601
600
599
598
597
596
595
594
593
592
591
590
589
588
587
586
585
584
583
582
581
580
579
578
577
576
575
574
573
572
571
570
569
568
567
566
565
564
563
562
561
560
559
558
557
556
555
554
553
552
551
550
549
548
547
546
545
544
543
542
541
540
539
538
537
536
535
534
533
532
531
530
529
528
527
526
525
524
523
522
521
520
519
518
517
516
515
514
513
512
511
510
509
508
507
506
505
504
503
502
501
500
499
498
497
496
495
494
493
492
491
490
489
488
487
486
485
484
483
482
481
480
479
478
477
476
475
474
473
472
471
470
469
468
467
466
465
464
463
462
461
460
459
458
457
456
455
454
453
452
451
450
449
448
447
446
445
444
443
442
441
440
439
438
437
436
435
434
433
432
431
430
429
428
427
426
425
424
423
422
421
420
419
418
417
416
415
414
413
412
411
410
409
408
407
406
405
404
403
402
401
400
399
398
397
396
395
394
393
392
391
390
389
388
387
386
385
384
383
382
381
380
379
378
377
376
375
374
373
372
371
370
369
368
367
366
365
364
363
362
361
360
359
358
357
356
355
354
353
352
351
350
349
348
347
346
345
344
343
342
341
340
339
338
337
336
335
334
333
332
331
330
329
328
327
326
325
324
323
322
321
320
319
318
317
316
315
314
313
312
311
310
309
308
307
306
305
304
303
302
301
300
299
298
297
296
295
294
293
292
291
290
289
288
287
286
285
284
283
282
281
280
279
278
277
276
275
274
273
272
271
270
269
268
267
266
265
264
263
262
261
260
259
258
257
256
255
254
253
252
251
250
249
248
247
246
245
244
243
242
241
240
239
238
237
236
235
234
233
232
231
230
229
228
227
226
225
224
223
222
221
220
219
218
217
216
215
214
213
212
211
210
209
208
207
206
205
204
203
202
201
200
199
198
197
196
195
194
193
192
191
190
189
188
187
186
185
184
183
182
181
180
179
178
177
176
175
174
173
172
171
170
169
168
167
166
165
164
163
162
161
160
159
158
157
156
155
154
153
152
151
150
149
148
147
146
145
144
143
142
141
140
139
138
137
136
135
134
133
132
131
130
129
128
127
126
125
124
123
122
121
120
119
118
117
116
115
114
113
112
111
110
109
108
107
106
105
104
103
102
101
100
99
98
97
96
95
94
93
92
91
90
89
88
87
86
85
84
83
82
81
80
79
78
77
76
75
74
73
72
71
70
69
68
67
66
65
64
63
62
61
60
59
58
57
56
55
54
53
52
51
50
49
48
47
46
45
44
43
42
41
40
39
38
37
36
35
34
33
32
31
30
29
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
1

FIGURE 92

MQLTRCCFVFLVQGSILVLVICQDDGPPGSEDPERDDHEGQPRPRVKRGHISPKSRPMANSTL
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVANLLVTGKIVDHGNGTFSVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMEWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWSCSQPKVVCVYIAFYSTDYRLVQKVCVDNYHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCCTCATTGCCCTCGGGCCTGCGCTGCC
CTTTATGTCTTACCATGCCATCGAGCCGGTGCATCATCTTCCTCATGCCGGAGCTTCCTT
CTGGTTGGTGTCTACTGATTCGTCCTGTTGGTCATGGCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAAATCTGCTGATCTTGAGCGTTGCTCTGTCTATATCCAAGAA
ATGTTCCGATTTGCATATTATAAACTCTTAAAAAAAGCCAGTGAAGGTTGAAGAGTATAAACCC
AGGTGAGACAGCACCCCTATGCGACTGCTGGCCTATGTTCTGGCTGGCTTGAATCATGA
GTGGAGTATTTCCCTTGTGAATACCCATCTGACTCCTGGGGCCAGGCACAGTGGCATTCA
GGAGATTCTCCTCAATTCTCCTTATTCACTGACGCTGGTCAATTATCTGCTGCATGT
ATTCTGGGCATTGTATTTTGATGGCTGTGAGAAGAAAAAGTGGGCATCCCTTATCGTC
TCCTGACCCACCTGCTGGTGTCAAGCCCAGACCTTCATAAGTTCTTATTATGGAATAAACCTGGCG
TCAGCATTATAATCCTGGTGCATGGGCACCTGGCATTCTAGCTGCGGGAGGCAGCTGCCG
AAGCCTGAAACTCTGCCTGCTTGCCAAGACAAGAACCTTCTTACAACCAGCGCTCCAGAT
AACCTCAGGGAACCAGCACCTCCAAACCGCAGACTACATCTTAGAGGAAGCACAACGTGCCT
TTTCTGAAAATCCCTTTCTGGTGAATTGAGAAAGAAATAAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAGPALALYVFTIAIEPLRIIFLIAGAFFWLVSILLISSLVWFMARVIIDNKDG
PTQKYLILFGAFVSVIQQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFMSGV
FSFVNTLSDSLGP GTVGIGHDSPQFFLYSAFMTLVII LLHVFWGIVFFDGCEKKWGILLIVLLT
HLLVSAQTFISSYYGINLASAFIILVLMGTWAAGGSCRSLKLCLLCQDKNFLLYNQRSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTCACCAGAGTAAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTGCCTCGTG
GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTGCTTAGGATCAAC
TCGGTCATTACCACAGCTAAACCTGCTTGGACTCCCTCCCACAAAACTGGCTCCGGATCAGG
GAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTCCTCTTAAGTCTGATAACCATTAACA
CAGATGCTCACACTGGGCCAGATCTGCATCTGTTAAATCCTGCTGCAGGAATGACACCTGGTAC
CCAGACCCACCCATTGACCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
CAATTTTGTCAACAAACTTGGAGGCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC
·TTCACGAGCCTCATCCATTCTGTTCCCGGGAGGCATCCTGCCACCAGTCAGGCAGGGC
TAATCCAGATGTCCAGGATGGAAGCCTCCAGCAGGAGGAGCAGGTGTAATCCTGCCACCCAGG
GAACCCCAGCAGGCCGCCTCCAACCTCCAGTGGCACAGATGACGACTTGCAGTGACCACCCCT
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGAAATTCA
GTAAGCTTTCAAATTTCAACTAAGCTGCCTCGAATTGGTGTACATGTGAATCTTATC
ATTGATTATATTGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAAATTCTTAATT
TACCTGAAAATATTCTGAAATTTCAGAAAATATGTTCTATGTAGAGAATCCCAACTTTAAAAAA
CAATAATTCAATGGATAAATCTGTCTTGGAAATATAACATTATGCTGCCTGGATGATATGCATAT
TAAAACATATTGGAAAAACTGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM
LTLGPDHLNLNPAAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

FIGURE 97

GCTCAAGTGCCTGCCCTGCCCCACCCAGCCAGCCTGGCAGAGCCCCCTGGAGAAGGAGCTCT
CTTCTGCTGGCAGCTGGACCAAGGGAGGCAAGTCTTGGCCTGGAGGGCTGTCTGACCATG
GTCCCTGCCCTGGCTGTGGCTGTTGTCTCCGTCCTCCAGGCTCTCCCAAGGCCAGCCTGC
AGAGCTGTCTGTGGAGTTCCAGAAAACATATGGTGGAAATTCCCTTATACCTGACCAAGTTC
CGCTGCCCTGTGAGGGGCTGAAGGCCAGATCGTGTCTGAGGGACTCAGGAAGGCAACTGAG
GGCCCATTGCTATGGATCCAGATTCTGGCTCCCTGCTGGTACCAAGGGCCCTGGACCGAGAGGA
GCAGGCAGAGTACCAAGCTACAGGTACCCCTGGAGATGCAGGATGGACATGTCTTGTTGGGTCAC
AGCCTGTCTGTGACCGTGAAGGATGAGAAATGACCAGGTGCCCATTTCTCAAGCCATCTAC
AGAGCTGGCTGAGCCGGGTACCAAGGCCATCCCTCTCTCCCTGAGGCTTCAGACCG
GGATGAGGCCAGCACAGCAACTCGGATCTCGATTCCACATCTGAGCCAGGCTCCAGGCCAGC
CTTCCCCAGACATGTCCAGCTGGAGCCTGGCTGGGCTCTGGCCTCAGCCCCAAGGGGAGC
ACCAGCCTTGACCAACGCCCTGGAGAGGACCTACAGCTTGTGACAGGTCAAGGACATGGGTGA
CCAGGCCCTCAGGCCACCAGGCCACTGCCACCGTGGAGTCTCCATCATAGAGGACCTGGGTGA
CCCTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGTCTTATACCCGACCATGGGCCAGGTA
CACTGGAGTGGGGGTATGTGCACTATCACCTGGAGGACCATCCCCCGGACCTTGAAGTGA
TGCAGAGGGAAACCTCACGTGACAGAGAGCTGGACAGAGAACGCCAGGCTGAGTACCTGCTCC
AGGTGCGGGCTCAGAATTCCATGGCAGGACTATGCGCCCTCTGGAGCTGCACGTGCTGGTGA
ATGGATGAGAAATGACAACAGTGCCTATGCCCTCCCCGTGACCCACAGTCAGCATCCCTGAGCT
CAGTCCACCAGGTACTGAAGTGAAGTGAAGACTGTCAGCAGGAGATGCAGATGCCCGGCTCCCCA
ATTCCCACTGTGTATCAGCTCTGAGCCCTGAGCCTGAGGATGGGTAGAGGGGAGACCTTC
CAGGTGGACCCCACTCAGGCACTGTGACGCTGGGGTCTCCACTCCGAGCAGGCCAGAACAT
CTGCTTCTGGTCTGGCATGGACCTGGCAGGGTGGCTCAGCAGCACGTGTAAG
TCGAAGTGCAGTCACAGATAATCAATGATCACGCCCTGAGTCTCATCACTTCCAGATTGGGCT
ATAAGCCTCCCTGAGGATGTGGAGCCCCGGACTCTGGTGGCCATGCTAACAGCATTGATGCTGA
CTCTGAGGCCCTTCCGCTCATGGATTTCCTGCAATGGAGAGGGGAGACACAGAGGGACTTTG
GCCTGGATTGGAGCCAGACTCTGGGCTATGGTAGACTCAGACTCTGCAAGAACCTCAGTTATGAG
GCAGCTCCAAGTCAAGGTTGGTGGTGGTGGCTGAGAGTGTGGCAGAGCTGGTGGGGCAGGGCC
AGGCCCTGGACCCGCCCCAGGTGACTGTGCTAGTGGAGAGAGTGTGATGCCACCCCCAAGTGG
ACCAAGGAGAGCTGAGGGCAGTGTCCCCATCAGTGCCTCAGCCGGCTTTCTGCTGACCATC
CAGCCCTCCGACCCCATCAGCCGAACCCCTCAGGTTCTCCCTAGTCATGACTCAGAGGGCTGGCT
CTGCTTGGGAGAAATTCTCCGGGAGGTGCACACCAGCTCCCTGAGGGCAGGCCAGGCTGGGG
ACACCTACAGGTGTTGTGGAGGCCCCAGGATAACAGCCTGACTCTGCCCTGTGCCCTCCCAA
TACCTCTGCACACCCGCCAAGACCATGGCTTGTGAGTGGACCCAGCAAGGACCCGATCT
GGCCAGTGGCACGGTCCCTACAGCTCACCTGGTCCAACCCCCACGGTGCACAGGGATTGGC
GCCTCCAGACTCTCAATGGTCCCATGCCCTACCTCACCTGGCCCTGCATTGGGGAGGCCAGT
GAACACATAATCCCCGTGGTGGTGCAGCCACAATGCCAGATGTGGCAGCTCTGGTGTGAGTGT
CGTGTGCTGCAACGTGGAGGGCAGTGCATGCGCAAGGTGGCCGATGAAGGGCATGCCA
CGAAGCTGTGGCAGTGGCATTCTGTAGGCACCCCTGGTAGCAATAGGAATCTCCTCATCCTC
ATTTCACCCACTGGACCATGTCAGGAAGAAGGACCCGGATCAACCAGCAGACAGCGTCCCCCT
GAAGGGCAGACTGTCTGAATGGCCCAAGGCAGCTAGCTGGAGCTGGCCTCTGGCTCCATCTGAG
TCCCCTGGGAGAGAGGCCAGCACCCAAAGATCCAGCAGGGACAGGACAGAGTAGAAGCCCTCCA
TCTGCCCTGGGGTGGAGGACCATCACCACATCACCAGGATGTCTGAGAGCTGGACACCAACT
TATGGACTGCCCATGGAGTGTCCAAATGTCAGGGTTGGCCATAATAAGCCAGAGAA
CTGGGCTGGGCCATGGAAAAAAGA

FIGURE 98

MVPAWLWLLCVSVPQALPKAQPAELSVEVPENYGGNFLYLTKLPLPREGAEQQIVLSGDGKAT
EGPFAMDPDSGFLLVTRALDREEQAELYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVPHFSQAI
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQSPDMFQLEPRLGALALSPKG
STS LDHALERTYQLLVQVKDMGDQASGHQATATVEVSIIESTWVSLPEIHLAENLKVLYPHHMAQ
VHWSSGGDVHYHLESHPGPFEVNAEGNLYVTRELDREAQAELYLLQVRAQNSHGEDYAAPPLELHVL
VMDENDNVPICPFRDPTVSIPELSPPGTEVTRLSAEDADAPGSPNSHVYQLLSPEPEDGVEGRA
FQVDPTSGSVTLGVPLRAGQNILLVLA MDLAGAEGGFSSTCEVEVA VTDINDHAPEFITSQIG
PISLPEDVEPGTLVAMLTAIDADLEPAFRLMDFAIERGDTEGTFG LDWEPSGHVRLRLCKNLSY
EAAPSHEVVVVVQSVAKLGVPGPGPGATATVTVLVERVMPPPQLDQESYEASVPISAPAGSFLLT
IQPSDPISRTLRFSLVNDSEGWL CIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALTLPVPS
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGPNPTVQRDWRLQTLNGSHAYLT LALHWEP
REHIIPVVVSHNAQM WQLLVRVIVCRCNVEGQC MRKVGRMKGMPTKLSAVGILVGTLVAIGIFLI
LIFTHWTMSRKDKPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCACGCCGT
AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCTGGAAAGCACCAGCCTTTA
TCTCTCACCTCAAGTCCCCTTCTCAAGAACATCCTCTGTTCTTGCCTCTAAAGTCTTGGTAC
ATCTAGGACCCAGGCATCTGCTTCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG
TTCTCTTATGTTGGTCAACTATTGCAATTAGAAGCTGCAACAAATTCCAATGAGACTAGCACC
TCTGCCAACACTGGATCCAGTGTGATCTCAGTGGAGCCAGCACAGCCACCAACTCTGGTCCAG
TGTGACCTCCAGTGGGTCAAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCAATGGGTCA
GCATAGTCACCAACTCTGAGTCCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
TTCAAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACCTCCAGTGG
GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCCTCAGTGGGCCAGCACAGTCACCAACT
CTGGTCCAGTGTGACCTCCAGTGGAGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCC
AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCAGTGGGCCAGCACAGCCAC
CAACTCTGACTCCAGCACACCTCCAGTGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCAGTCAGTGGGCCAGCACAGTGG
GCCACCAACTCTGAGTCCAGCACACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAG
AACGACCTCCAATGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGCCA
GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGCCAGCACTGCCACCAACTCTGAG
TCCAGCACGACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG
GGCTAGCACAGCCACCAACTCTGACTCCAGCACAACCTCCAGTGGGCCAGCACAGCCACCAACT
CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACACCCCTCC
AGTGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGCCAACACAGCCAC
CAACTCTGAGTCCAGCACAGTGTCCAGTGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAA
CCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCTCCAGTGGGCCAGCAC
GCCACCAACTCTGACTCCAGCACACCTCCAGTGGGCCAGCACAGCCACCAACTCTGAGTCTAG
CACAGTGTCCAGTGGGATCAGCACAGTCACCAATTCTGAGTCCAGCACACCTCCAGTGGGCCA
ACACAGCCACCAACTCTGGGCCAGTGTGACCTCTGAGGCTCTGGAACAGCACAGCTCTGACTGGA
ATGCACACAAACTTCCCATAGTCATCTACTGCACTGAGTGGCAAGCCTGGTGGGCTCTGGT
GCCGTGGGAAATCTTCCATCACCCCTGGCTCGGTTGTGGGCCGTGGGCTCTTGTCTGGC
TCTTCTCTGTGAGAAACAGCCCTGCTGAGAAACACCTTTAACACAGCTGTCTACCAACCT
CATGGCTCAACCATGGCCCTGGTCCAGGCCCTGGAGGGAAATCATGGAGCCCCCACAGGCCAG
GTGGAGTCCTAACCTGGTTCTGGAGGAGACCACTGATCATCGATAGCCATGGAGATGAGCGGGAGGA
ACAGCGGGCCCTTGAGCAGCCCCGGAAAGCAAGTGCCTGAGTCAGGAGAAGAGACCTGGCA
CCCAAGACCTGGTTCTTCTCATTCATCCAGGAGACCCCTCCAGCTTGTGAGATCTGAA
AATCTGAGAAGGTATTCCCTCACCTTCTTGCCTTACAGACACTGGAAAGAGAAATACTATAT
TGCTCATTTAGCTAAGAAAATAACATCTCATCTAACACACAGCAAAGAGAAAGCTGTGCTTG
CCCCGGGGTGGGTATCTAGCTGAGATGAACACTCAGTTAGGAGAAAACCTCCATGCTGGACTC
CATCTGGCATTCAAATCTCCACAGTAAAATCCAAAGACCTCAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MKMQKGNVLLMFGLLHLLEAATNSNETSTSANTGSSVISSGASTATNKGSSVTSSGVSTATISGS
SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTSGASTATNSESSTPSS
GASTVTNKGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNDSSTTSSGASTA
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTSNGAGTATNSES
STTSSGASTATNDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNDSSTTSS
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
TNSESSTTSSGVSTATNSESSTTSSGASTATNDSSTSSEASTATNSESSTVSSGISTVTNSES
STTSSGANTATNKGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGLVPWEIFLITLVVVA
AVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGPGPGENHGAPHRPRWSPNWFWRPVSSI
AMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCCTCCGGTTACGGGATGAATTAAACGGCGGGTTCGCACGGAGGGTGTGACCCCTA
CGGAGCCCCAGCTTGCACGCCACTCGCGTCGCGCGCGTGCCTGCTGTACAGGTG
GGAGGCTGGAACTATCAGGCTAAAAACAGAGTGGTACTCTCTGGAAAGCTGGCAACAAAT
GGATGATGTGATATGCAITCCAGGGAAAGGGAAATTGTGGTGCTTCTGAACCCATGGTCAATT
AACGAGGCAGTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCTTGGAAATCAT
GGTGTCATGGAAAGGGATTACTTATACTGACTCTGGGGAGCTTTGGAAAGCATT
TCATGCTGAGTCCCTTTTACCTTGATGTTGTAACCCATCTGGTATCGCTGGATCAACAC
CGCCTTGCAACATGGCTACCCCTACCTGTGGCATATTGGAGACCAGTTGGTGTAAAAGT
GATTATAACTGGGATGCATTGTTCTGGAGAAAGAAGTGTCAATTATCATGAACCACCGGACAA
GAATGGACTGGATGTTCTGTGGAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT
TGCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTGGTTGGGCATGCAGGCTGCTGCCTATAT
CTTCATTCAAGGAAATGGAAGGATGACAAGAGCCATTCGAAGACATGATTGATTACTTTGTG
ATATTCA~~G~~CAACCACCAACTCAACTCCTCATATTCCCAGAAGGGACTGATCTCACAGAAAACAGCAAG
TCTCGAAGTAATGCATTGCTGAAAAAAATGGACTTCAGAAATATGAATATGTTTACATCCAAG
AACTACAGGCTTACTTTGTGGTAGACCGTCAAGAGAAGGTAAGAACCTTGATGCTGTCCATG
ATATCACTGTGGCGTATCCTCACAAACATTCTCAATCAGAGAAGCACCTCCTCCAAGGAGACTT
CCCAGGGAAATCCACTTCACGTCCACCGGTATCCAATAGACACCCCTCCCCACATCCAAGGAGGA
CCTTCAACTCTGGGCCACAAACGGTGGGAAGAGAAAGAAGAGAGGCTGCGTCCATATCAAG
GGGAGAAGAATTTTATTTACCGGACAGAGTGTCAATTCCACCTTGAAGTCTGAACGACTGGTC
CTTGTGGTCAAATTGCTCTATACTGTATTGGACCCGTCAAGGCTGCAATGTGCCTACTCAT
ATATTGTACAGTCTGTTAAGTGGTATTATAATCACCATTGTAATCTTGTGCTGCAAGAGA
GAATATTGGTGGACTGGAGATCATAGAACATTGCATGTTACCGACTTTACACAAACAGCCACAT
TTAAATTCAAAGAAAATGAGTAAGATTAAAGGTTGCCATGTGAAAACCTAGAGCATATTG
GAAATGTTCTAAACCTTCTAAGCTCAGATGCATTGCACTGACTATGCGAATATTCTTACT
GCCATCATTATTGTTAAAGATATTGCACTTAATTGTGGAAAAAATATTGCTACAATT
TTAATCTCTGAATGTAATTGATACTGTGTACATAGCAGGGAGTGATGGGTGAAATAACTT
GGGCCAGAATATTAAACAAATCATCAGGCTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSFGIMVSWKGIYFILTLFWGSFFGSIFMLSP
FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFPGERSVIIIMNHRTRMDWM
FLWNCLMRYSYLRLEKICLKASLGVPFGWAMQAAAYIFIHRKWDDKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSRSNAFAEKNGLQKYEVVLHPRTTGFTVVDRREGKNLDLVHDITVA
YPHNIPQSEKHLLQGDFPREIHFHVRYPIDLPTSKEDLQLWCHKRWEERLRSFYQGEKNF
YFTGQSVIPPCKSELRLVVKLLSILYWTLFSPAMCLLIYSLVKWYFIITIVIFVLQERIFGG
LEIIIELACYRLLHKQPHLNNSKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGACGGGCTCGAGTGAAGAGCCTCTCACGGCTCTGCCTGAGACAGCTGGCCTGACC
TCCAAATCATCCATCCACCCCTGCTGTCACTGTTTCATAGTGTGAGATCAACCCACAGGAATA
TCCATGGCTTTGTGCTCATTTGGTTCTCAGTTCTACAGAGCTGGTGTCAAGGACAGTGGCAAGT
CACTGGACCGGGCAAGTTGTCCAGGCCCTGGTGGGGAGGACGCCGTGTTCTCCTGCTCCCTCT
TTCTGAGACCACTGAGGGCTATGGAAGTGCGGTTCTCAGGAATCAGTTCCATGCTGTGGTC
CACCTCTACAGAGATGGGAAGACTGGGAATCTAACAGATGCCACAGTATCGAGGGAGAACTGA
GTTTGTGAAGGACTCCATTGCAAGGGGGCGTGTCTCTAAGGCTAAAAAACATCACTCCCTCGG
ACATGCCCTGTATGGGTGCTGGTCAGTCCCAGATTACGATGAGGAGGCCACCTGGGAGCTG
CGGGTGGCAGCACTGGCTCACTTCCATCGTGGATATGTTGACGGAGGTATCCA
GTTACTCTGCCTGCTCAGGTGGTCCCCAGCCCACAGCCAAGTGGAAAGGTCCACAAGGAC
AGGATTGTCTTCAGACTCCAGAGCAAATGCAGATGGTACAGCTGTATGATGTGGAGATCTCC
ATTATAGTCCAGGAAAATGCTGGAGCATATTGTGTTCCATCCACCTTGCTGAGCAGAGTCATGA
GGTGGAAATCCAAGGTATTGATAGGAGAGACGTTTCCAGCCCTCACCTGGCGCTGGCTTCTA
TTTACTCGGGTACTCTGTGGTGCCTGTGTGGTTGTCATGGGATGATAATTGTTCTTC
AAATCCAAAGGAAAATCCAGCGGAACGGACTGGACTGGAGAAGAACGGACAGGACAGGAGAATTGAG
AGACGCCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTACCCGAAGCTCTGCG
TTTCTGATCTGAAAATGTAACCCATAGAAAAGCTCCCAGGGAGGTGCCTCACTCTGAGAAGAGA
TTTACAAGGAAGAGTGTGGTGCCTCTCAGGGTTTCAAGCAGGGAGACATTACTGGGAGGTGGA
CGTGGGACAAAATGAGGGTGTATGTGGAGGTGTCGGGATGACGTAGACAGGGGGAAAGAAC
ATGTGACTTTGTCTCCAACAATGGGTATTGGGTCTCAGACTGACAACAGAACATTGTATTTC
ACATTCAATCCCCATTATCAGCCTCCCCCCCAGCACCCCTCACAGCAGTAGGGGTCTTCC
GGACTATGAGGGTGGGACCATCTCTCTCAATAAAATGACCACTCCCTTATTATACCCCTGC
TGACATGTCAAGTGTGAGGCTGTGAGACCCATATCCAGCATGGATGTGACGAGGAAAAG
GGGACTCCCATATTCAATGTCAGTGTCTGGGATGAGACAGAGAACCCCTGCTTAAAGGGC
CCCACACCAGACAGGAGACACAGCCAAGGGAGGTGCTCCCGACAGGTGGCCCAGCTTCC
CCGGAGCCTGGCACAGAGTCAGCCCCCAGTCAGCTCTTCTAGGGAGCTGAGGTTCTCTGCC
TGAGCCCTGCAGCGGCAGTCACAGCTCCAGATGAGGGGGATTGGCTGACCCCTGTGGAG
TCAGAAGCCATGGCTGCCCTGAAGTGGGACGGAATAGACTCACATTAGGTTAGTTGTGAAAA
CTCCATCCAGCTAAGCGATCTGAACAAGTCACAACCTCCCAGGCTCCTCATTGCTAGTCACGG
ACAGTGATTCTGCCCTCACAGGTGAAGATAAAGAGACAAACGAATGTGAATCATGCTGCAGGTT
TGAGGGCACAGTGTGCTAATGATGTTTATATTACATTTCACCCATAAACTCTGTT
TGCTTATTCCACATTAATTACTTTCTATACCAAATCACCCATGGAATAGTTATTGAACACC
TGCTTGTGAGGCTCAAAGAATAAAGAGGGAGGTAGGATTTCACTGATTCTATAAGCCAGCAT
TACCTGATACCAAAACCAAGGCAAAGAAAACAGAAGAAGAGGAAGGAAAACACAGGTCCATATCC
CTCATTAACACAGACACAAAAATTCTAAATAAAATTAAACAAATTAAACTAAACAATATATTAA
AAGATGATATATAACTACTCAGTGTGGTTGTCCCACAAATGCAGAGTGGTTAATATTAAAT
ATCAACCAGTGAATTCAAGCACATTAAGTAAAAAGAAAACCATAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETS A E A M E V R F F R N Q F H A V V H
LYRDGEDWESKQMPQYRGRTEFKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
VAALGSIPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPGQGDLS SDRS RANADGYSLYDVEISI
IVQENAGSILCSIH LAEQSHEVESKVVLIGETFFQPSPWRLASILLGLLCGALCGVVVMGMIIVFFK
SKGKIQAELDWRRKHGQAEI RDARKHAVEVTLDPETAHPKLCVSDLKTVTHR KAPQEVPHSEKRF
TRKSVVASQGFQAGRHYWEVDVGQNVGVYVGCRDDVRGKNNVTLSPNNGYWVRLTTEHLYFT
FNPHFISLPPSTPPTRGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLLRPYIQHAMYDEEKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTCATTGCTGGTTGGCAATGATGATCGGCCAGATGTGGTGAGGGCTAGGAAAAGAG
TTTGTGGAAACCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGTCCGGCAGTGTGCATTGGA
CTCACTGTTCAATTATGTGAGATATAATCAAAGAAGACCTACAATTACTATAGCACATTGTCAATTACAAC
TGACAAACTATATGCTGAGTTGGCAGAGAGGCTTAACAAATTACAGAAATGAGCCAGAGACTTGAAT
CAATGGTGAAGGATTCATTTTATAAATCTCCATTAAGGGAAAGAATTGTCAAGTCTCAGGTTATCAAGTTC
AGTCAACAGAACGATGGAGTGTGGCTCATATGCTGTTGATTGAGATTCACTCTACTGAGGATCCTGA
AACTGTAGATAAAATTGTTCAACTGTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCCCTAAAGTAG
ATCCTCACTCAGTAAATTAAAAAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTGCTGCGGA
ACACGAAGAAGTAAACTCTAGGTCAAGGTCTCAGGATGTTGGGGACAGAAGTGAAGAGGGTGAATG
GCCCTGGCAGGCTAGGCTGCAGTGGGATGGGAGTCATCGCTGTGGAGCAACCTTAATTATGCCACATGGC
TTGTGAGTGCTGCTCACTGTTTACAACATATAAGAACCCCTGCCAGATGACTGCTTCCTTGGAGTAACA
ATAAAACCTTCGAAATGAAACGGGTCTCCGGAGAATAATTGTCATGAAAATACAAACACCCATCACA
TGACTATGATATTCTCTTGAGAGCTTCTAGCCCTGTTCCCTACACAAATGCACTGACATAGAGTTGTC
TCCCTGATGCATCCTATGAGTTCAACCAGGTGATGTGATGTTGTGACAGGATTGGAGCACTGAAAAT
GATGGTTACAGTCAAATCATCTCGACAAGCACAGGTGACTCTCATAGCGCTACAACTTGCAATGAACC
TCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGCTGGCTCCCTAGAAGGAAAACAGATGCAT
GCCAGGGTGAECTGGAGGACCCTGGTTAGTTCAAGATGCTAGAGATATCTGGTACCTGCTGGAATAGTG
AGCTGGGAGATGAATGTGCGAACCCAAACAAGCCTGGTGTAACTAGAGTTACGGCCTTGCAGGACTG
GATTACTTCAAAACTGGTATCTAAGAGACAAAAGCCTCATGGAACAGATAACATTGTTGTTTGTGTTTG
GGTGTGGAGGCCATTAGAGATACAGAATTGGAGAAGACTGCAAAACAGCTAGATTGACTGATCTCA
ATAAAACTGTTGTTGATGCTGATTTCTCCAGCTCTGTCAGCTCGCACGTAAGCATTGCTTCGCTG
GATCAACTCTGTCTGAGCAATAGTTGAAACATTATGTACATAGAGAAAATAGATAATAACATATTAC
ATTACAGCCTGTATTCAATTGTTCTAGAAGTTGTCAGAATTGACTGTCAGATTTGACATAAAATTGTAAT
GCATATATAACATTGAAAGCACTCCTTTCTTCAGTTCCCTCAGCTCTCATTCAAGCAAATATCCATT
TCAAGGTGAGAACAAAGGACTGAAAGAAAATATAAGAAGAAAATCCCTACATTGTTATTGGCACAGAA
AAAGTATTAGGTGTTTCTTAGGAAATTAGAAAATGATCATATTCAATTGAAAGGTCAAGCAAAGACA
GCAGAACATCAACTTCATCAATTAGGAAGTATGGGAACTAAGTTAAGGAAGTCCAGAAAGAAGCCAAG
ATATATCCTTATTTCATTCCAAACAAACTACTATGATAAAATGTGAAGAAGATTCTGTTTTGTGACCT
ATAATAATTACAAACTTCATGCAATGTAATTGTTCTAAGCAAATTAAAGCAAATTAAACATTG
TTACTGAGGATGTCAACATATAACAATAAAATATAACACCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVI FISLIVLAVCIGLT VHYVRYNQKKTYNYYSTLSFTTDKLY
AEFGREASNNFTEMSQRLESMVKNAFYKSPLREEFVKSQVIKFSQQKHGVLAHMLLICRFHSTED
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG
TEVEEGEWPWQASLQWDGSHRCGATLINATWLVAHHCFTTYKNPARWTASFGVTIKPSKMKGRL
RRIIVHEKYKHP SHDYDISLAELSSPV PYTNAVHRVCLPDASYEFQPGDVMFVTGFGALKNDGYS
QNHLRQAQVTLIDATT CNEPQAYNDAITPRMLCAGSLEGKTDACQGDGGPLVSSDARDI WYLAG
IVSWGDEC A KPNKPGVYTRVTALRDWITSKTGI

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAACCGTCTCCAGCTGAAGCCAATGCAGCCCCCTCCGGCTCTCCGCAAGAAGTCCCTG
CCCCGATGAGCCCCCGCGTGCCTCCCCGACTATCCCAGGCGGGCTGGGGCACCGGGCCAGC
GCCGACGATCGCTGCCGTTGGCCCTGGGAGTAGGATGTGGTAAAGGATGGGCTTCCT
ACGGGGCTACAATGCCAGAGAAGATTGGTGAAGTGTCTGCCTGCTTGATGAGGGACTACCTAA
TCTGCTCTTGGTAATGCCATCAGTGTGGAGTTCTGCTTGATGAGGGACTACCTAA
ATAATGTTCTCACTTAACGCAGAACAGGGTAGAGGAAGCAGTCATTGACTTACCTTCCT
GTGGTTCATCCGGTCACTGATTGCTGTTGCTTCCTTACATTGTTGGGATGTTAGGATATTG
TGGAACGGTAAAAGAAATCTGTTGCTTCTGCTGGACATATGAACAGGAACCTATGGTTCCAGTACAATGGTCA
GTGAGAACTGGCTGTGGCTTGGACATATGAACAGGAACCTATGGTTCCAGTACAATGGTCA
GATATGGTCATTTGAAAGCCAGGATGACAAATTGATTACCTAGATATGGCTTACTCA
TGCTTGAATTTTTCAGAGAGAGTTAAGTGTGGAGTAGTATATTCACTGACTGGTTGG
AAATGACAGAGATGGACTGGCCCCCAGATTCTGCTGTGTTAGAGAATTCCAGGATGTTCCAAA
CAGGCCACCAGGAAGATCTCAGTGACCTTATCAAGAGGGTGTTGGAGAAAATGTTACCTT
TTTGAGAGGAACCAAACACTGCAGGTGCTGAGGTTCTGGAAATCTCATTGGGTGACACAAA
TCCTGCCATGATTCTCACCAATTACTCTGCTCTGGCTCTGTATTATGATAGAAGGGAGCTGG
ACAGACCAATGATGTCCTGAAGAATGACAACCTCTCAGCACCTGTCATGTCCTCAGTAGAACT
GTTGAAACCAAGCCGTCAAGAATCTTGAAACACACATCCATGGCAAACAGCTTAATACACACT
TTGAGATGGAGGAGTTA~~AAAA~~AGAAATGTCACAGAAGAAAACCACAAACTTGTGTTATTGACT
TGTGAATTTGAGTACATACTATGTGTTGAGAAATATGAGAAATAAAAATGTTGCCATAAAA
TAACACCTAACGATATACTATTCTATGCTTAAATGAGGATGGAAAAGTTCATGTCATAAGTC
ACCACCTGGACAATAATTGATGCCCTTAAATGCTGAAGACAGATGTCATACCCACTGTTAGCC
TGTGTATGACTTTACTGAACACAGTTGAGGAGCTGAGGTTGATTAGCATTTCGCA
TCCATGCAAACGAGTCACATATGGGGACTGGAGCCATAGTAAGGTTGATTACTCTACCAA
CTAGTATATAAGACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTATTA
CTCAGCGATCTATTCTGCTGCTAAATAATTATATACAGAAAACCTTCATATTGGTACT
ACCTAAATGTTGATTTGCTGTTACTAAAATATCTTACCACTAAAAGAGCAAGCTAACACAT
TGTCTTAAGCTGATCAGGGATTTTGCTGTTACTAAAATATCTTACCACTAAAAGAGCAAGCTAACACAT
TTCAGTTCTGATAATGTTAAGAATAACCAATTGAAAGAAAATTGTCCTGATAGCATCATT
ATTTTACGCTTCCCTGTTAAAGCTTACTATTCTGCTGGCTTATATTACACATATAAC
TGTATTTAAATACTAACCAATTGAAAGAAAATTACCACTGTTGATAGCATAGGAATCATTATTC
AGAATGTTGAGTCTGCTTTAGAAGTTAAGAAAATTGCACTAACACTTGTGATTCTGAG
AAGGACTTGTATGCTGTTCTCCAAATGAAAGACTCTTTGACACTAAACACTTTTAAAGAAA
GCTTATCTTGCCCTCTCCAAACAAGAAGCAATAGTCCTGCAAGTCAATATAAAATTCTACAGAAAA
TAGTGTCTTTCTCCAGAAAATGCTGTTGAGAATCATTAAACATGTGACAATTAGAGATT
CTTGTGTTATTCACTGATTAATATACTGTTGAAATTACACAGATTATTAAATTGTTACAA
GAGTATAGTATATTATTGAAATGGAAAAGTGCATTACTGTTGATTTGTGATTGTTAT
TTCTCAGAATATGAAAGAAAATTAAATGTCATAAAATATTCTAGAGAGTAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTLTAETRVEEAVILTYFPVVHP
VMIAVCCFLIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMVT
LKARMNTNYGLPRYRWLTHAWNFFQREFKCCGVVFIDWLEMTEMDWPPDSCCVREFPGCSKQAHQ
EDLSDLYQEGCGKKMYSFLRGTKQLQVLRLGISINGTQILAMILTLLWALYYDRREPQTDQM
MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGGACACCTTATCCCACCTCATCCTCATCCTCTTCCTGATAAAGCCCTACCAGTGCT
GATAAAAGCTTTCTCGTGAGAGCCTAGAGGCCTTAAAAAAAAGCTTGAAGAGAAGGGACAAAGGAACA
CCAGTATAAGAGGATTTCAGTGTTCAGTGGCAGTTGGCAGAAGGATGCTCCATTCCCTGCTTCTCACCTG
CCTCTCATCACAGGCACCTCCGTCAACCCGTGGCCCTAGATCCTGTTCTGCTTACATCAGCCTGAATGAGC
CCTGGAGGAACACTGACCACCACTGGATGAGTCACAGGTCCCTCTATGTGACAACCATGTGAATGGGAG
TGGTACCACTTCACGGCATGGGGAGATGCCATGCCTACCTTCTGCATACAGAAAACCACTGTGAAACCCA
CGCACCTGCTGGCTAACAGCACCCCCCTAGAAGGCGACGGCATTGTGCAACGCCAGGCTTGTGCCAGCT
TCAATGGAACACTGCTGTCTGGAAACACCACGGTGGAAAGTCAAGGCTTGCCTGGAGGCTACTATGTGATCGT
CTGACCAAGCCCAGCGCTGCTTCCACGTACTGTGGTCAATTATGACATCTGCGACGAGGACTGCCATGG
CAGCTGCTCAGATAACCAACCGAGTGACATGCGCTCCAGGAACGTGCTAGGCCCTGACAGGCAGACATGCTTG
ATGAAAATGAATGTGAGCAAACAAACGGTGGCTGCAAGAAGTGAAGACTGTGAAGACCTCAAAACTCCTACCGCTGT
GAGTGTGGGTTGGCGTGTGCTAAGAAGTGAAGACTGTGAAGACGTTGAAGGATGCCACAATAACAA
TGGTGGCTGCAGCCACTCTTGCCTGGATCTGAGAAAGCTACCGAGTGTGAATGTCCCCGGGCCTGGTGT
CTGAGGATAACCACACTTGCCTGGCTGAGATCTGAGGAAACCTCCCTGCGAGGAGTGTCAACGGCACCCATGTCAACAT
CTGGTTGGTGGCCTGGAGCTCTTCCCTGACCAACACCTCCCTGCGAGGAGTGTCAACGGCACCCATGTCAACAT
CCTCTCTCTCAAGACATGTGGTACAGTGGTCAAGTGTGGTGAATGACAACATTGTGGCCAGCAACCTCGTGA
CAGGTCTACCCAAGCAGACCCGGGACAGCGGGACTTCATCATCCGAACCAAGCTGCTGATCCGGTG
ACCTGCGAGTTCCACGCCGTACACCATTCTGAAGGATACTGTTCCCAACCTTCGAAACTCCCACTGGAAT
CATGAGGCCAAATCATGGGATCTTCCCACTCTGGAGATCTCAAGGACAATGAGTTGAAGAGCCTTAC
GGGAAGCTCTGCCACCCCTCAAGCTTGTGACTCCCTCTACTTTGGCATTGAGCCCGTGGTGCACGTGAGCGGC
TTGGAAAGCTTGGTGGAGAGCTGCTTGCCTACCCCAACCTCCAAGATCGACGAGGCTCTGAAATACTACCTCAT
CCGGGATGGCTGTGTTCACTGACTCGGTAAGCAGTACACATCCGGGATCACCTAGCAAAGCACTTCCAGG
TCCCTGCTTCAAGTTGTGGGCAAAGACCCACAAGGAAGTGTCTGCACTGCCGGTTCTGTGTGAGTG
TTGGACGAGCGTCCCGCTGTGCCACGGGTTGCCACCGGCAATCGTCGGAGGAGGACTCAGC
CGGTCTACAGGGCCAGACGCTAACAGCGGCCGATCCGCATCGACTGGGAGGACTAGTTGTAGCCATACCTC
GAGTCCCTGCATTGGACGGCTCTGCTTTGGAGCTTCTCCCCCACGCCCTCAAGAACATCTGCCAACAGC
TGGGTTCACTGAGTTCAACTGAGCTCAGACTCCCAGCACCAACTCACTGATTCTGGTCCATTCACTGGGCA
CAGGTCACTGCTGAACAATGTGGCCTGGGTTGGGTTCTGTTCTAGGGTTGAAAACAAACTGTCCA
CCCAGAAAGACACTCACCCATTCCCTCATTTCTTCAACTTAAACACTCGTGTATGGTGAATCAGAC
CACAAATCAGAAGCTGGGTATAATATTCAAGTTACAAACCTAGAAAATTAAACAGTTACTGAAATTATGA
CTTAAATACCCAATGACTCCTTAAATATGTAATTATAGTTACCTGAAATTCAATTCAAATGCAGACTAA
TTATAGGAATTGGAGTGTATCAATAAAACAGTATATAATT

FIGURE 110

MPPFLLTCLFITGTSVSPVALDPCSAYISLNEPWRNTDHQLDSEQGPPLCDNHVNGEWYHFTGMAGDAMP
TFCIPENHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCCLWNNTTVEVKACPGGYYVRLTKPSVCFH
YCGHFYDICDEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGGCSEICVNLKNSYRCECGVGRV
LRSDGKTCEDVEGCHNNNNGCShSCLGSEKGYQCECPRLVLSEDNHTCQVPVLCKSNAIEVNIPRELVGG
LELFLTNTSCRGVSNGLHVNLFLSKTCGTVVVDVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLIPV
CEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFITLEIFKDNEFEEPYREALPTLKLRSLSLYFGIEPVVHV
SGLESLVESCFATPTSKIDEVLKYYLIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKEVGKDHKEVFLHCRV
LVCGVLDERSRCAQGCHRRMRRGAGGEDSAGLQQQLTGGPIRIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCCTGCACTCGG
GCCTCCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCAGCCAGGACCTGTGTGGGAGGCCCT
CCTGCTGCCTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGATGT
TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAAACCCCGTATCCCC
ATGGAGACCTTCAGAAAGGTGGGATCCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATATTGT
GGTTGTCCATCAAGGTGATTCTGGATAAATACTACTTCCTCTGCCGGCAGCCTCTCACTTCATCCGA
GGAAGCAGCTGTGACGGAGAGCTGGACTGTCCTGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC
CCCGAAGGGCCTGCAGTGGCAGTCCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTGCCAC
AGGGAACTGGTTCTGCCTGTTGACAACTTCACAGAAGCTCGCTGAGACAGCCTGTAGGCAGATGG
GCTACAGCAGAGCTGTGGAGATTGGCCAGACAGGATCTGGATGTTGAAATCACAGAAAACAGCCAG
GAGCTTCGCATGCGAACACTCAAGTGGCCCTGTCCTCAGGCTCCCTGGCTCCCTGCACTGTCTGCCCTG
TGGGAAGAGCCTGAAGACCCCCCGTGGTGGTGGGGAGGAGCCTGTGATTCTGGCCTTGGCAGG
TCAGCATCCAGTACACAAACAGCACAGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCTCACGGCA
GCCCACTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAAACTGGCAG
CTTCCCACCCCTGGCTGTGGCCAAGATCATCATCATTGAATTCAACCCCATGTACCCAAAGACAATGACA
TCGCCCTCATGAAGCTGCAGTTCCACTCATTCTCAGGCACAGTCAGGCCCATCTGCTGCCCTCTT
GATGAGGAGCTCACTCCAGCCACCCACTCTGGATCATTGATGGGCTTACGAAGCAGAATGGAGGGAA
GATGTCGACATACTGTCAGGGCTCAGTCAGGCTATTGACAGCACACGGTCAATGCAGACGATGCGT
ACCAGGGGAAGTCACCGAGAAGATGATGTGTCAGGCATCCCGAAGGGGTGTGGACACCTGCCAGGGT
GACAGTGGTGGGCCCTGATGTACCAATCTGACCAGTGGCATGTGGTGGCATGTTAGCTGGGCTATGG
CTGCGGGGGCCCGAGCACCCAGGAGTATAACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCT
GGAAGGCTGAGCTGTAATGCTGCTGCCCTTGCAGTGCTGGAGGCCCTCCCTGCCACCTGCCAC
GGGATCCCCAAAGTCAGACACAGAGCAAGAGTCCCTGGTACACCCCTCTGCCACAGCCTCAGCAT
TTCTGGAGCAGCAAAGGGCTCAATTCCCTGTAAGAGACCCCTCGCAGCCAGAGGCGCCAGAGGAAGTCA
GCAGCCCTAGCTGGCCACACTTGGTGCTCCAGCATCCAGGGAGGAGACACAGCCACTGAACAAGGTCT
CAGGGTATTGCTAAGCCAAGAAGGAACCTTCCACACTACTGAATGGAAGCAGGCTGTCTGTAAAGCC
CAGATCACTGTGGCTGGAGAGGAGAAGGAAAGGTCTGCCAGCCCTGTCCGTCTTACCCATCCCCAA
GCCTACTAGAGCAAGAAACCAAGTGTAAATATAAAATGCACTGCCCTACTGTGGTATGACTACCGTTACCT
ACTGTGTCAATTGTTATTACAGCTATGCCACTATTATAAGAGCTGTAACTCTCTGGCAAAAAAAA
AAAA

FIGURE 112

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPIIIALLSLASIIIVVVLIKVILDKYFLCG
QPLHFI PRKQLCDGELDCPLGEDEEHCVKS FPEGPAVAVRLSKDRSTLQVLD SATGNWFSACFDN
FTEALAE TACRQM GYSRAVEIGPDQ DLDV VEITEN SQEL RMRN SGPCLSGSLVSLHCLACGKSL
KTPRVVGEEASVDSWPWQVSIQYDKQHVCGGSILDPHWVLTAAHCFRKHTDVFNWKVRAGSDKL
GSFPSLAVAKIIIIEFNPMPYPKNDI ALMKLQFPLTFSGT VRPICL PFFDEELTPATPLWIIGWG
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDT CQGDGGPLMYQS
DQWHVVGVIVSWG YGC GGPSTPGVYTKVSAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACCTCTGGTCCCAGTGATCCACCGCCTAGCCTCCAAGGTGCTGTGATTA
TAGGTGTAAGCCACCGTGTGGCCTCTGAACAACTTTCAGCAACTAAAAAGCCACAGGAGT
TGAACTGCTAGGATTCTGACTATGCTGTGGCTAGTGCTCCTACTCCTACCTACATTAAATC
TGTTTTTGTCTTGTAACTAGCCTTACCTCTAACACAGAGGATCTGTCACTGTGGCTCT
GGCCCAAACCTGACCTTACCTGGAACGAGAACAGAGGTTCTACCCACACCGTCCCGAAG
CCGGGGACAGCCTCACCTGCTGGCCTCTCGCTGGAGCAGTGCCCTACCAACTGTCTCACGTCT
GGAGGCACTGACTCGGGCAGTGCAGGTAGCTGAGCCTTGGTAGCTGCCCTCAAGGTGGC
CTTGCCCTGGCCGTAGAAGGGATTTGACAAGCCCAGATTCATAGGCATGGCTCCACTGCC
AGGCATCAGCCTGCTGTAGTCATCAACTGCCCCGGGCGCAGGACGGCCGTGGACACCTGCTCA
GAAGCAGTGGGTGAGACATCACGCTGCCGCCATCTAACCTTTCATGTCCTGCACATCACCTG
ATCCATGGCTAACTGAACTCTGCTTCAAGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC
CAGAAGGGGCTGCTTAGACCACCTGGTTATGTGACAGGACTTGCAATTCTCTGGAACATGAGG
GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAACTTGCCAAATTATGGGTAGAAAAGATG
GAGGTGTTGGTTATACAAGGCATCGAGTCTCCTGCATTCACTGGACATGTGGGGAAAGGGCTG
CCGATGGCGATGACACACTCGGGACTCACCTCTGGGCCATCAGACAGCCGTTCCGCCCGAT
CCACGTACCAAGCTGCTGAAGGGCAACTGCAGGCCATGCTCATCAGCCAGGCAGCAGCCAAA
TCTGCGATCACCAGCCAGGGCAGCCGTGGGAAGGAGCAAGCAAAGTGACCATTTCTCTCCC
CTCCTCCCTCTGAGAGGCCCTCTATGTCCTACTAAAGCCACCAGCAAGACATAGCTGACAGG
GGCTAATGGCTCAGTGTGGCCAGGAGGTAGCAAGGCCTGAGAGCTGATCAGAAGGGCTGCT
GTGCGAACACGGAAATGCCCTCAGTAAGCACAGGCTGCAAAATCCCAGGCAAAGGACTGTG
CTCAATTAAATCATGTTCTAGTAATTGGAGCTGTCCTAACAGACAAAGGAGCTAGAGCTGGTT
CAAATGATCTCCAAGGGCCCTATACCCCAGGAGACTTGATTTGAATTGAAACCCCAAATCCA
AACCTAAGAACCAAGGTGCATTAAGAACATCAGTTATTGCCGGGTGTGGTGGCCTGTAATGCCAACAT
TTTGGGAGGCCAGGGCGAGGCGGGTAGATCACCTGAGGTAGCAAGACCAGCCTGCCAACATGG
TGAAACCCCTGTCCTACTAAAAATACAAAAAAACTAGCCAGGCATGGTGGTGTGCCCTGATC
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACTGGAGGTGAAGGAGGCTGAGACA
GGAGAATCACTCAGCCTGAGCAACACAGCAGACTCTGTCTCAGAAAAAATAAAAAAGAATT
TGGTTATTTGTAA

FIGURE 114

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLAELTRAVQVAEPLGSCGFQGGPCPGRRD

Signal peptide:

amino acids 1-15

FIGURE 115

CAGCAGTGGCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA
AGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTAAATCCAAGAAA
ATATGTAATCACTTAAGATTGTGGACTGGTGTGTTGCTCCTGCCCTAACTCTAATTGTCCT
GTTTGGGGAGCAAGCACTCTGGCCGGAGCTACCCAAAAAGCCTATGACATGGAGCACACTT
TCTACAGCAATGGAGAGAAGAAGAAGATTACATGGAAATTGATCCTGTGACCAGAACTGAAATA
TTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTAAAAACGGATACACTGG
CATCTACTCGTGGGTCTCAAAAATGTTTATCAAACACTCAGATTAAAGTGAATTCTGAATTTT
CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCAACTTCTTGAAACAGTCAGTG
ATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTCTTAAAAATTCCAAAATTCTGGA
GATTGTGATAACGTGACCAGTATTGGATCAATCCCACCTCTAATATCAGTTCTGAGTTACAAG
ACTTTGAGGAGGAGGGAGAAGATCTCACTTCTGCAACGAAAAAAAAGGGATTGAACAAAAT
GAACAGTGGGTGGTCCCTCAAGTGAAGTAGAGAAGACCCGTACGCCAGACAAGCAAGTGAGGA
AGAACTTCCAATAATGACTATACTGAAATGGAATAGAATTGATCCCAGTGTGAGAGAG
GTTATTGTTGATTACTGCCGTGAGGCAACCGCTATTGCCGCCGTCTGTGAAACCTTACTA
GGCTACTACCCATATCCATACTGCTACCAAGGAGGACGAGTCATCTGCGTGTGTCATGCCCTG
TAACTGGTGGGTGCCCGCATGCTGGGGAGGGTCTAATAGGAGGTTGAGCTCAAATGCTAAAC
TGCTGGCAACATATAATAATGCATGCTATTCAATGAATTCTGCCATGAGGCATCTGGCCCT
GGTAGCCAGCTCCAGAATTACTGTAGGTAATTCTCTCTCATGTTCTAATAAAACTTCTACA
TTATCACCAAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFSKKICKSLKICGLVFGILALTLLIVLFWGSKHF
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLEVHDFKNGYTG
EFSEPEEEIDENEEITTFEQSVIWVPAEKPIENRDFLKN
SKILEICDNVTMYWINPTLISVSE
LQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQ
ASEEELPINDYTENGIEFDPM
GRV
ERGYCCIYCRRGNRYCRRVCEPLLGYYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-
242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCTCAGGAGCGCGTTAGCTTCACACCTTGGCAGCAGGAGGGCGGCAGCTTCTGCAGGGCGCA
GGCGGGCGGCCAGGATCATGTCCACCACCATGCCAAGTGGTGGCGTCTCTGTCCATCCTGGGCT
GGCGGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGACAACCCCGTCACCT
CCGTGTCCAGTACGAAGGGCTCTGGAGGAGCTCGTGAGGCAGAGTTCAAGGCTTCACCGAATGCAGGCC
TATTCACCACCTGGACTTCCAGCATGCTGCAGGCAGTGCAGGAGCCCTGATGATCGTAGGCATCGCCT
GGGTGCCATTGGCCTCTGGTATCCATCTTGGCCTGAAATGCATCCGATTGGCAGCAGGGACTCTG
CCAAAGCCAACATGACACTGACCTCGGGATCATGTTCAATTGTCAGGTCTTGTGCAATTGCTGGAGTG
TCTGTGTTGCCAACATGCTGGTACTAACTTCTGGATGTCACAGCTAACATGTACACCGGATGGTGG
GATGGTGCAGACTGTTCAAGACCAGGTACACATTGGTGCAGGCTGTTGGCTGGCTGGCGTGGAGGCC
TCACACTAACATTGGGGTGTGATGATGTCATGCCCTGCCGGGCTGGCACCAAGAAAACCAACTACAAA
GCCGTTCTTATCATGCCCTAGGCCACAGTGTGCCTACAAGCCTGGAGGCTCAAGGCCAGCACTGGCTT
TGGGTCCAACACCAAAACAAGAAGATAACGATGGAGGTGCCGACAGAGGACGAGGTACAATCTTATC
CTTCCAAGCACGACTATGTGTAATGCTCAAGACCTCTCAGCACGGCGGAAGAAAACCCGGAGAGCTCA
CCCCAAAAACAAGGAGATCCCATCTAGATTCTTCTGCTTTGACTCACAGCTGGAAAGTTAGAAAAGCCT
CGATTCATCTTGAGAGGCCAAATGGCTTAGCCTCAGTCTGTCTAAATATCCACCATAAAACA
GCTGAGTTATTATGAATTAGAGGCTATAGCTCACATTCAATCCTCTATTCTTTTAAATATAACT
TTCTACTCTGATGAGAGAATGTGGTTAATCTCTCTCACATTGATGATTAGACAGACTCCCCCTC
TTCCCTCTAGTCATAAAACCCATTGATGATCTATTCCAGCTTACCCAGAAAACCTTGTGAAAGGAAA
GAGTAGACCCAAAGATGTTATTCTGCTTTGAAATTGTCTCCCCACCCCAACTGGCTAGTAATAA
ACACTTACTGAAGAAGCAATAAGAGAAAAGATATTGTAATCTCCAGCCATGATCTGGTTTCTT
ACACTGTGATCTAAAGTACCAACCAAGTCATTTCAGTTGAGGCAACCAACCTTCTACTGCTG
TTGACATCTTCTATTACGACAACCAATTCTAGAGTTCTGAGCTCCACTGGAGTCCTCTTCTGT
CGCGGGTCAGAAATTGTCCTAGATGAATGAGAAAATTATTTTTAAATTAGCTAAATAGTTAA
AATAAAATAATGTTAGTAAATGATACTATCTGTGAAATGCTCACCCCTACATGTGGATAGAAG
GAAATGAAAAAAATAATTGCTTGACATTGTCTATATGGTACTTTGTAAGTCAAGTCAAGTACAAATTCC
ATGAAAAGCTCACACCTGTAATCTAGCACTTGGAGGCTGAGGAGGAAGGATCACTTGAGGCCAGAAGT
TCGAGACTAGCCTGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAATCAGCCAGTCA
TGGTGGCATAACCTGTAGTCCAGCATTCCGGAGGCTGAGGTGGAGGATCACTTGAGGCCAGGGAGGT
TGGGGCTGCAGTGAGCCATGATCACACCACTGCACCCAGGTGACATAGCGAGATCCTGTCTAAAAA
AATAAAAATAATAATGGAACACAGCAAGTCCTAGGAAGTAGGTTAAACTAATTCTTTAA

FIGURE 118

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVFQYEGLWRSCVRQSSGFTECRP
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA
CRGLAPEETNYKAVSYHASGHHSVAYKPGGFKA
STGFGSNTKNKKIYDGGARTEDEVQSYP SKHDY
V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

© 2002 University of Michigan Health System

FIGURE 119

GGAAAAACTGTTCTTCTGTGGCACAGAGAACCCGTCAAAGCAGAAGTAGCAGTCGGAGTCC
AGCTGGCTAAACATCCCAGAGGATAATGGCAACCCATGCCTTAGAAATCGCTGGCTGTTCTTG
GTGGTGTGGAATGGTGGCACAGTGGCTGTCAGTGTCTGCCTCAGTGGAGAGTGTGGCCTTCATT
GAAAACAACATCGTGGTTTGAAAACCTCTGGAAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA
CATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTTCTCCGGACCTACAGGCAGCCAGAG
GAUTGATGTGCTGCTTCCGTATGTCCTCTGGCTTCATGATGCCATCCTGGCATGAAATGC
ACCAGGTGCACGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTCAT
CATCACGGGCATGGTGGTGCATCCCTGTGAGCTGGTGCAATGCCATCATCAGAGATTCTATA
ACTCAATAGTGAATGTTGCCAAAAACGTGAGCTGGAGAAGCTCTACTTAGGATGGACCACGGCA
CTGGTGTGATTGTTGGAGGAGCTGTTCTGCGTTTTGTTGCAACGAAAAGAGCAGTAGCTA
CAGATACTCGATACCTCCCATCGCACACCCAAAAAGTTACACACCGGAAAGAAGTCACCGAGCG
TCTACTCCAGAAGTCAGTATGTGTAGTTGTATGTTTAACTATAAGCCATGCAAATG
ACAAAAATCTATATTACTTCTCAAATGGACCCAAAGAAACTTGTGATTTACTGTTCTTAACTGCCT
AATCTTAATTACAGGAACGTGCATCAGCTATTGATTCTATAAGCTATTCAGCAGAATGAGATA
TTAAACCCAATGTTGATTGTTCTAGAAAGTATGTAATTGTTCTAAGGTGGTCAAGCATCTA
CTCTTTTATCATTACTCTAAATGACATTGCTAAAGACTGCATTATTTACTACTGTAATTCTCC
ACGACATAGCATTATGTACATAGATGAGTGTAAACATTATCTCACATAGAGACATGCTTATGGT
TTTATTAAAATGAAATGCCAGTCCATTACACTGAATAATAGAACTCAACTATTGCTTTCAAGGGAA
ATCATGGATAGGGTGAAGAAGGTTACTATTAATTGTTAAAAACAGCTAGGGATTAATGCTTCCA
TTTATAATGAAGATTAAAATGAAGGCTTAATCAGCATTGTAAGGAAATTGAATGGCTTCTGATAT
GCTGTTTTAGCCTAGGAGTTAGAAATCTAACTTCTTATCCTCTCCAGAGGCTTTTT
CTTGTGTATTAAATTAAACATTAAAACGCAGATATTGTCAGGGCTTGCAAGGGCTTGCAACTGCTT
TTCCAGGGCTATACTCAGAAGAAGATAAAAGTGTGATCTAAGAAAAAGTGTGTTAGGAAAGTG
AAAATTTTGTGTTGTATTGAAAGAAGATGATGCATTGACAAGAAATCATATATGTATGGAT
ATATTAAATAAGTATTGAGTACAGACTTGTGAGGTTCATCAATATAAAAGAGCAGAAAATA
TGTCTGGTTTCATTGCTTACAAAAACAAACAAAAAAAGTTGTCTTGTGAGAACTTCACCT
GCTCCTATGTGGGTACCTGAGTCATTGTCATTGTTCTGTGAAAAATAAAATTCTTCTGTA
CCATTCTGTTAGTTTACTAAAATCTGAAATACTGTATTCTGTTATTCAAATTGATGAA
ACTGACAATCCAATTGAAAGTTGTGTCGACGTCTGCTAGCTAAATGAATGTGTTCTATTGCTT
TATACATTATTAATAAAATTGTACATTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTVAVTVMPOWRVSIFIENNIIVFENFWEGLWMNCVRQANIRMQCK
IYDSLALSPDLQAARGLMCAASVMSFLAFMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG
MVVLIPVSWVANAIIRDIFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCCNEKSSY
RYSIPSHRTTQKSHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

© 2007 Pearson Education, Inc.

FIGURE 121

GGAGAGAGGCCGCGCGGGTGAAGGCCATTGATGCAGCCTGCGGCGCCTCGGAGCGCGGAG
CCAGACGCTGACCACGTTCTCCTCGCTCCCTCCAGCTCCGCGCTGCCCGCAGCC
GGGAGCCATGCGACCCAGGGCCCCGCCGCCTCCCGCAGCGGCTCCGCGGCCCTGCTGCTCC
TGCTGCTGAGCTGCCCGGCCGTGAGCGCCTCTGAGATCCCCAAGGGAAAGCAAAAGGCCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC
TGGTCGAGACGGGAGCCCTGGGCAATGTTATTCCGGTACACCTGGATCCCAGGTGGATG
GATTCAAAGGAGAAAAGGGGAATGTCTGAGGGAAAGCTTGAGGAGTCCTGGACACCCAACACTAC
AAGCAGTGTTCATGGAGTTCAATTGAATTATGGCATAGATCTGGAAAATTGCGGAGTGTACATT
TACAAAGATGCGTTCAAATAGTGTCTAACAGAGTTTGTTCAGTGGCTCACTCGGCTAAATGCA
GAAATGCATGCTGTCAGCGTTGGTATTCACATTCAATGGAGCTGAATGTTCAGGACCTTCCC
ATTGAAGCTATAATTATGGACCAAGGAAGCCCTGAAATGAATTCAACAATTAAATTATCATCG
CACTTCTCTGTGGAAGGACTTGTGAAGGAATTGGGCTGGATTAGTGGATGTTGCTATCTGGG
TTGGCACTTGTTCAGATTACCCAAAAGGAGATGCTTCACTGGATGGAATTCAAGTTCTCGCATC
ATTATTGAAGAACTACCAAATAAATGCTTAATTTCATTTGCTACCTCTTTTATTATGCC
TTGGAATGGTCACTAAATGACATTAAATAAGTTATGTATACATCTGAATGAAAAGCAAAG
CTAAATATGTTACAGACCAAAAGTGTGATTCACACTGTTAAATCTAGCATTATTCAATTG
CTTCAATCAAAGTGGTTCAATTAGTTGTTAGAATACTTCTCATAGTCACATT
CTCTCAACCTATAATTGGAATTGTTGTTCTTTGTTCTTAGTATAGCATTAA
AAAAAATATAAAAGCTACCAATTGTACAATTGTAATGTTAAGAATTGTTATATCTGT
TAAATAAAATTATTCCAACA

FIGURE 122

MRPQGPAAASPQRRLRGLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAGVPGR
DGSPGANVIPGTPGIPGRDGFKGEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK
MRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECGPLPIEAIYLDQGSPEMNSTINIHTS
SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCCGGTACGGGCTCTCTGCCTCTGGCTCCAACGCAGCTCTGTGGCTGAA
CTGGGTGCTCATCACGGAACTGCTGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
ATTGCTGGAAGAATACATCAIGTTTCGATAAGAAGAAATTGTAAGGATCCAGTTTTTTTA
ACCGCCCCCTCCCCACCCCCAAAAACTGAAAGATGCAAAAACGTAATATCCATGAAGATCC
TATTACCTAGGAAGATTTGATGTTGCTGCGAATGCCGTGTTGGGATTATTGTTCTGGAG
TGTCTCGTGGCTGGCAAAGAATAATGTCAAAATCGGCCATCTCCAAGGGGTCCTTAA
TCTTCTGGGTGTCAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGGTCATGCAACTG
GCCCTAAGCAAAGCAAAGACCTAAGGACGACCTTGAACAATACAAAGGATGGGTTCAATG
TAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTATAGCCCCACTGTCTACTGACAATG
CTTCTCTGCCAACGAGGATGCCCTAAGGGCTGTAGGTGTGAAGGCAAATGGTATATTGTA
ATCTCAGAAATTACAGGAGATAACCCCTCAAGTATATCTGCTGGTTGCTTAGGTTGTCCTCGCT
ATAACAGCCTTAAAAACTTAAGTATAATCAATTAAAGGCTAACAGCTCACCTGGCTATAC
CTTGACCATAACCATACTAGCAATATTGACGAAATGTTTAATGGAATACGAGACTCAAAGA
GCTGATTCTTAGTCCAATAGAATCTCCTATTCTTAACAATACCTCAGACCTGTGACAAATT
TACGAACTTGGATCTGCTCTATAATCAGCTGCATTCTCTGGGATCTGAACAGTTGGGCTTG
CGGAAGCTGCTGAGTTACATTACGGCTAACCTCCCTGAGAACCATCCCTGTGCGAATATCCA
AGACTGCCAACCTGAAACTTGGACCTGGATATAACCGGATCCGAGTTAGCCAGGAATG
TCTTGCTGGCATGACTCAAAGAACCTCACCTGGAGCACAATCAATTTCAGCTCAAC
CTGGCCCTTTTCCAAGGTTGGTCAGCCTCAGAACCTTACTTGCAAGTGGAAATAAAATCAGTGT
CATAGGACAGACCATGTCCTGGACCTGGAGCTCCTAACAAAGGCTGATTATCAGGCAATGAGA
TCGAAGCTTCAGTGGACCCAGTGTGTTCCAGTGTGTCGGCAATCTGCAGCGCTCACCTGGAT
TCCAACAAGCTCACATTATTGGTCAAGAGATTGGATTCTGGATATCCCTCAATGACATCAG
TCTTGCTGGGAATATGGGAATGCAGCAGAAATATTGCTCCCTGTAACCTGGCTGAAAGTT
TTAAAGGTCTAAGGGAGAATACAATTATCTGTGCGAACAGCTGCAAGGAGTAAATGTG
ATCGATGAGACTACAGCATCTGGCAAAAGTACTACAGAGGTTGATCTGGCAG
GGCTCTCCAAAGCCGACGTTAACGGCAAGCTCCCCAGGCCGAAGCATGAGAGCAAACCCCTT
TGCCCCCGACGGTGGAGCCACAGAGCCCCCCCAGAGACCGATGCTGACGCCAGCACATCT
TTCCATAAAATCATCGCGGGCAGCGTGGCGCTTCCCTGTCGTGTCATCTGCTGGTTAT
CTACGTGTCATGAAAGCGGTACCCCTGCGAGCATGAAGCAGCTGCAAGCGCTCCCTCATGGAA
GGCACAGGAAAAGAACAGTCCCTAACGCAAATGACTCCCAGCACCCAGGAATTATGTA
GATTATAAACCCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGGACGGGACCCCTGCACCTA
TAACAAATCGGGCTCCAGGGAGTGTGAGGTATGAACCTGGTGAATGAGCTCTAAAGCT
GGGAATAAGTGGTGTCTTATGAACTCTGGTGAATCAAGGGAACGCGATGCCCTCCCT
TTCCCTCTCCCTCACTTGGTGGCAAGATCCTCCCTGTGCGTTAGTGCATTATAACT
GGTCAATTCTCTCATACATAATCAACCCATTGAAATTAAATACCACAAATCAATGTGAAGCTT
GAACCTGGTTAATATAACCTATTGATAAGACCCCTTACTGATTCCATTATGTCGCATT
GTTTAAGATAAAACTTCTTCATAGGTAAAAAAAAAA

100 90 80 70 60 50 40 30 20 10

FIGURE 124

MGFNVIRLLSGSAVALVIAPTVLLTMSSAERGCPKGCRCEGKMWYCESQKLOEIPSSISAGCLG
LSLRYNSLQKLKYNQFKGLNQLTWLYLDHNHISNIDENAFNGIRRLKELILSSNRISYFLNNTR
PVTNLRLNLLSYNQLHSLGSEQFRGLRKLLSLHLRSNSLRTIPVRIFQDCRNLELLDLGYNRIRS
LARNVAGMIRLKELHLEHNQFSKLNLAFLPRLVSLQNLQWNKISVIGQTMSTWSSLQRSDL
SGNEIEAFSGPSVFQCVPNLQRNLDSNKLFIGQEILDWSWLNDISLAGNIWECSRNICSLVN
WLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVILLVIYVSWKRYPASMKQLOQR
SLMRRHRKKKRQSLKQMTPSTQE FYVDYKPTNTETSEMLLNGTGPCTYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTCGGCTACTGCTGAATGTCCGTCCCCGGAGGAGGAGGGCTTTGCCGCTG
ACCCAGAGATGGCCCCGAGCGAGCAAATTCTACTGTCCGGCTGCCGGCTACCGTGGCCGAGCT
AGCAACCTTCCCTGGATCTCACAAAAACTGACTCAAATGCAAGGAGAACGCAGCTCTGCTC
GGTTGGGAGACGGTGCAAGAGAACTGCCCCCTATAGGGAAATGGTGCACAGCCCTAGGGATC
ATTGAAGAGGAAGGCTTCTAAAGCTTGGCAAGGAGTGACACCCGCCATTTACAGACACGTAGT
GTATTCTGGAGGTCGAATGGTACATATGAACATCTCCGAGAGGTTGTGTTGGCAAAAGTGAAG
ATGAGCATTATCCCTTGAAATCAGTCATTGGAGGGATGATGGCTGGTGTATTGGCCAGTT
TTAGCCAATCCAATGACCTAGTGAAGGTTAGATGCAAATGAAAGGAAAAGGAAACTGGAAGG
AAAACCATTGCGATTCGTGGTGTACATCATGCATTGCAAAATCTTAGCTGAAGGAGGAATAC
GAGGGCTTGGCAGGCTGGTACCCAATATAACAAAGAGCAGCAGCAGCTGGTAATATGGGAGATT
ACCACTTATGATACTGAAACACTACTTGGTATTGAATACACCAGTGGAGGACAATATCATGAC
TCACGGTTTATCAAGTTATGTTCTGGACTGGTAGCTCTATTCTGGAACACCAGCCGATGTCA
TCAAAAGCAGAATAATGAATCAACCACGAGATAAACAGGAAGGGACTTTGTATAAAATCATCG
ACTGACTGCTTGATTCAAGGTGAAGGATTCATGAGTCTATATAAAGGCTTTTAC
ATCTGGCTGAGAATGACCCCTGGTCAATGGTGTCTGGCTTACTTATGAAAAAATCAGAGAGA
TGAGTGGAGTCAGTCCATTTAA

FIGURE 126

MSVPEEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMQGEAALARLGDGARES
APYRGMVRTALGIIEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHLREVVFGKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMOMEGRKRKLEGKPLRFRGVHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCGLVASILGTPADVIKSRI MNQP
RDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 127

CGCGGATCGGACCCAAAGCAGGT CGGC GGCG CAGGAGAGCGGCCGGCGTCAGCT CCTCGAC
CCCCGTGTCGGGCTAGTCCAGCGAGGCGGACGGCGCGTGGGCCATGGCCAGGCCGGCATGG
AGCGGTGGCGCGACCGGCTGGCGCTGGTGACGGGGGCTCGGGGGCATCGGC GCGCCGTGGCC
CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGC GCCGACTGTGGCAACATCGAGGA
GCTGGCTGCTGAATGTAAGAGTGCAAGGCTACCCC GGACTTTGATCCC TACAGATGTGACCTAT
CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTAGACATC
TGCATCAACAATGCTGGCTTGGCCGGCCTGACACCCCTGCTCTCAGGCAGCACCCAGTGGTTGGAA
GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGA
AGGAGCGGAATGTGGACGATGGCACATCATTAACATCAATAGCATGTCAGGTGGCCACCGAGTGT
CCCCCTGTCGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTACTGCCGTGACAGAGGGACT
GAGGCAAGAGCTTGGGAGGCCAGACCCACATCCGAGCCACGTGCATCTCCAGGTGTGGTGG
AGACACAAATTGCCCTCAAACCTCCACGACAAGGACCTGAGAAGGCAGCTGCCACCTATGACCAA
ATGAAGTGTCTCAAACCGAGGATGTGGCGAGGCTGTTATCTACGTCCCTCAGCACCCCCGCACA
CATCCAGATTGGAGACATCCAGATGAGGCCACGGAGCAGGTGACTAGTGACTGTGGGAGCTCC
TCCTTCCCTCCCCACCCCTCATGGCTGCCCTGCCTCTGGATT TAGGTGTTGATTCTGGAT
CACGGGATACCACCTCTGTCCACACCCCCGACCAGGGCTAGAAAATTGTTGAGATTTTATA
TCATCTGTCAAATTGCTTCAGTTGTAATGTGAAAAATGGCTGGGAAAGGAGGTGGTGTCCC
TAATTGTTTACTGTTAACCTGTTCTGTGCCCCCTGGCACTTGGCTTGTCTGCTCTCAGTG
TCTTCCCTTGACATGGGAAAGGAGTTGTGGCAAATCCCCATCTTGTGACCTCAACGTCTG
TGGCTCAGGGCTGGGTGGCAGAGGGAGGCCCTCACCTTATATCTGTGTTATCCAGGGCTCC
AGACTTCCCTCTGCCCTGCCCCACTGCACCCCTCTCCCCCTATCTATCTCCTCTCGGCTCCC
AGCCCAGTCTGGCTTGTCCCCCTCTGGGT CATCCCTCACTGACTCTGACTATGGCAG
CAGAACACCAGGGCTGGCCAGTGGATTTCATGGT GATCATTAAAAAGAAAAATCGCAACCAA
AAAAAAAAAA

FIGURE 128

MARPGMERWRDRRLALVTGASGGIGAAVARALVQQGLKVVGCARTVGНИEELAAECKSAGYPGTLI
PYRCDLSNEEDILSMFSAIRSQHSGVDICINNAGLARPDTLLSGSTSGWKGDMFNVLALSICTR
EAYQSMKERNVDDGHIININMSGHRLPLSVTHFYSATKYAVTALTEGLRQELREAQTHIRATC
ISPQGVETQFAFKLHDKDPEKAATYEQMCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AACTTCTACATGGGCCTCCTGCTGCTGGTGCCTTCCTCAGCCCTCCTGCCGGTGGCCTACACCAT
CATGTCCCTCCCACCCTCCTTGACTGCGGGCCGTTCAAGGTGCAGAGTCAGTTGCCGGGAGC
ACCTCCCCCTCCGAGGCAGTCTGCTCAGAGGGCCTGGCCCAGAATTCCAGTTGGTTCATGC
CAGCCTGTAAAAGGCATGGAACCTGGGTGAATCACCGATGCCATTAAAGAGGGTTCTGCCA
GGATGGAAATGTTAGTCGTTCTGTCTGCCTGTTCAATTCAAGGCCACCTGTGG
CCGTTGAGTGCTTGAAATGAGGAACTGAGAAAATTAAATTCTCATGTATTTCATTTATTAA
TTAATTAACTGATAGTTGTACATATTGGGGGTACATGTGATATTGGATACTGTATA
TATATAATGATCAAATCAGGGTAACTGGGATATCCATCACATCAAACATTATTTTATTCTT
TTAGACAGAGTCACACTGTCACCCAGGCTGGAGTGCAGTGGTGCCATCTCAGCTTACTGCAAC
CTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCCTCACCTCCAAGTAGCTGGACTACAGGCAT
GCACCACAATGCCCAACTAATTGTATTAGTAGAGACGGGGTTTGCCATGTTGCCAGG
CTGGCCTGAACTCTGGCCTAAACAATCCACTTGCCCTGGCCTCCAAAGTGTATGATTACA
GGCGTGAGGCCACCGTGCCTGGCTAAACATTATCTTTCTTGTTGGAACTTGAAATTAT
ACAATGAATTATTGTTAACACTGTCATCTCCCTGCTGTGCTATGAAACACTGGACTCTCCCTCT
ATCTAACGTATATTGTACCAAGTTAACCAACCGTACTTCATCCCCACTCCTCTATCCTTCCC
AACCTCTGATCACCTCATTCTACTCTCACCTCCATGAGATCCACTTTAGCTCCACATGTG
AGTAAGAAAATGCAATATTGTCTTCTGTGCCCTGGCTTATTCACTTAACATAATGACTTCTG
TTCCATCCATGTTGCTGCAAATGACAGGATTGCTTAATTCAATTAAAATAACCACACATG
GCAAAAA

FIGURE 130

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSAREHLPSPRGSSLRGPRPRIPVLVSCQPV
KGHGTLGESPMFKRVFCQDGNVRSCVCVCAVFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTGTATAAAGACCTAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
ATCTTCCTCATCGGGACTAAAATTGGGCTGTTCTCAAGTAGCACCTCTATCAGTTATGGCTAAATCTG
TCCATCTGTGTGCGCTGCGATGCGGGTTCTTACTGTAATGATCGCTTCTGACATCCATTCAAACAG
GAATACCAGAGGATGCTACAACTCTACCTTGAGAACAAATAATAATGCTGGGATTCTTCAGAT
TTGAAAAACTGCTGAAAGTAGAAAGAATATACTTACCAACAGTTAGATGAATTCTACCAACCT
CCCAAAGTATGTAACAGAGTTACATTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTCAA
AAATTCCATCTGGAGAAATTACATTTAGATGACAACCTCTGTCTCTGCAGTTAGCATAGARGAGGGAGCA
TTCCGAGACAGCAACTATCTCGACTGCTTCTGTCCCGTAATCACCTAGCACAATTCCCTGGGTTT
GCCCAAGGACTATAGAAGAACTACGCTGGATGATAATCGCATATCCACTATTTCATCACCACCTCTTCAAG
GTCTCACTAGTCTAAACGCCCTGGTCTAGATGAAACCTGTGAACAAATCATGGTTAGGTGACAAAGTT
TTCTCAACCTAGTTAATTGACAGAGCTGCTCCCTGGTGCAGAATTCCCTGACTGCTGCACCAGTAAACCT
TCCAGGCACAAACCTGAGGAAGCTTATCTCAAGATAACCACATCAATCGGGTGCACCCCAAATGCTTTT
CTTATCTAAGGCAGCTATCGACTGGATATGCTCAATAAACTTAAGCTAATTACCTCAGGTATCTT
GATGATTGGACAATATAACACAACGTGATTCTCGCAACAATCCCTGGTATTGCGGGTGCAGATGAAATG
GGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACGTGCCTGGCTCATGTGCCAAGCCCCAGAAA
AGGTTCTGGGATGGCTATTAAGGATCTCAATGAGAACACTGTTGATGTAAGGACAGTGGGATTGTAAGC
ACCATTCAAGATAACCACGTCAATACCCAAACACAGTGTATCCGCCAAGGACAGTGGCCAGCTCCAGTGAC
CAAACAGCCAGATATTAAGAACCCCAAGCTCAACTAAGGATCAACAAACCACAGGGAGTCCCTCAAGAAAAA
CAATTACAATTACTGTGAAGTCTGTACCTCTGATACCATTCTATCTTGGAAACTTGCTCTACCTATG
ACTGCTTGAGACTCAGCTGGCTAAACTGGGCATAGCCCGCATTGGATCTATAACAGAAACAATTGT
AACAGGGAAACGCACTGGTACAGTGGTACAGCCCTGGAGCCTGATTCCCTATAAAGTATGCATGGTTC
CCATGGAAACCCAGCAACCTCTACCTATTGATGAAACCTCTGGTTGATGAGACTGAAACTGCACCCCTT
CGAATGTACAACCCATCAACCACCTCAATCGAGAGCAAGAGAAAGAACCTACAAAAACCCAAATTAC
TTTGGCTGCCATCTGGTGGGCTGTGGCCCTGGTACCTGCCCCTCTGCTTGTGTTAGGTGTTGGTATG
TTCATAGGAATGGATCGCTCTCTCAAGGAACCTGTGCATATAGCAAAGGGAGGAGAAGAAAGGATGACTAT
GCAGAAGCTGGCACTAAGAAGGACAACCTATCTGGAAATCAGGAAACTCTTTCAGATGTTACCAAT
AAGCAATGAACCCATCTCGAAGGAGGAGTTGTAATACACACCATTCTCTCTAATGGAATGAATCTGT
ACAAAAACAAATCACAGTGAAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC
TCACACTCATGATGCTGAAGGACTCACAGCAGACTGTGTTGGTTTAAACCTAAGGGAGGTGATG
GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRCDAFGFIYCNDRFLTSIPTGIPEDATTIYL
QNNQINNAGIPSIDLKNLLKVERIYLYHNSLDEFPTNLPKVKEHLQENNIRTITYDSLSKIPYL
EELHLDNSVS AVSIEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL
QGLTSLKRLVLDGNLLNNHGLDKVFFNLVNLTELSLVRNSLTAAPVNLPGTNLRKLYLQDNHIN
RVPPNAFSYLRQLYRLDMSNNNLSNLPQGIFDDLDNITQLILRNNPWYCGCKMKWVRDWLQSLPV
KVNVRGGLMCQAPEKVRGMAIKDLNAELFDCKDGSIVSTIQTITTAIPNTVYPAQGQWPAPVTKQPD
IKNPKLTQDQTTGSPSRKTITITVKSVTSDTIHISWKLALPMTALRLSWLKLGHSPAFGSITET
IVTGERSEYLVTALEPDSPYKVCMVPMETSNL LFDETPVCIETETAPLRMYNPTTLNREQEKE
PYKNPNLPLAAIIGGAVALVTIALLALVCWYVHRNGSLFSRNCAYSKGRRRKDDYAEAGTKKD NS
ILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNLYKNNHSESSSNRSYRDSGIPDSDHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCCTCCAGAGTCCTTGCCCAGGCCACCCAGGCTTCTGGCA
GCCCTGCCGGGCCACTTGTCTTCATGTCTGCCAGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG
TGCAGAGGCAGTCTGGCTGGCCAGAGCTCAGGGTGTGAGCGTGTGACCAGCAGTGAGCAGAG
GCCGCCATGCCAGCCTGGGCTGCTGCTCTGCTTAAGCTGACAGCACTGCCACCGCTGTGGT
CCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTCGAGACCTGATCCTGTCT
GCGCTGGAGAGAGCCACCGTCTCCTAGAACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT
GGGGGTCCGAGTGTGGAAGAGCAGCTAAAAAGTGTCCCCGAGAAGTGGGCCAGGAGGCCCTG
TGCAGCCGCTGAGCCTGCGCGTGGGATGCTGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
CTCCACTACCTCAAGCTGAGTGTACCTAACAGAGGTTCCAGTGCACCCCTCAGCCCG
GTTTGGAAGCTCCACATGCCTGGATCCACACTGATGCCCTTGGTGTACCCACGTTGGC
CCCAGGACTATTCTCAGAGGAGAGAAGTGCAGTGTGCCCTGGTGCAGCTGCTGGAAACCGGGACG
GACAGCAGCAGGCCCTGCCCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCGGCTGCTC
AGGCTACTGCCTGTCCCACCAACTGCTCTTCTCTGGGCCAGAATGAGGGGATGCACACAGG
GACCACTCCAACAGAGCCAGGACTATATCAACCTCTGCGCCAACATGATGGACTTGAACCGC
AGAGCTGAGGCCATCGGATACGCCAACCTACCCGGGACATCTCATGGAAAACATCATGTTG
TGGAATGGCGGCTTCTCCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA
AACAGCAGGAAGGATGCTCGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
TATCAGCAGCATTTTCGAGGAGAGTGAAGAGGGCGAGAAAAACATTCCAGATTCTCGCTCTG
TGCTCAGGCTGGAGTACAGTGGCGCAATCTGGCTACTGCAACCTTGCCCTGGTTCAAGC
AATTCTTGCCTCATCCTCCGAGTAGCTGGACTACAGGAGCGGCCACCATCTGGCTAAT
TTTTATATTTTTAGTAGAGACAGGGTTCATCATGTTGCTCATGCTGGTCTCGAACCTCTGAT
CTCAAGAGATCCGCCACCTCAGGCTCCAAAGTGTGGATTTAGGTGAGCCACCGTGTCTG
GCTGAAAAGCACTTCAAAGAGACTGTGTTGAATAAAGGGCCAAGGTTCTGCCACCCAGCACTC
ATGGGGCTCTCCCTAGATGGCTGCTCCCTCCACAAACAGCCACAGCAGTGGCAGGCCCTGG
GTGGCTTCTATACATCCTGGCAGAATACCCCCCAGCAAACAGAGAGGCCACACCCATCCACACCG
CCACCACCAAGCAGCCGCTGAGACGGACGGTCCATGCCAGCTGCCCTGGAGGAGGAACAGACCC
TTTAGTCCTCATCCCTTAGATCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG
ATAAGCAAAGCCACCCGACACCCAACTTGGAGCCAGGCTGAGTAGGCAGGGCAGGGTAGGTGG
GCCGGGAGGGACCCAGGTGTGAACGGATGAATAAAGTCAACTGCAACTGAAAAAAAAAAA

FIGURE 134

MSARGRWEGGGRRACRGSGLARAQGAERVTSSQRPMASLGLLLLLTLPLWSSSLPGLD
TAESKATIADLILSALERATVFLQEQLPEINLDGMVGVRVLEEQLKSVREKWAQEPLLQPLSLRV
GMLGEKLEAAIQRSIHYLKLSDPKYLREFQLTLQPGFWKLPHAWHTDASLVYPTFGPQDSFSEE
RSDVCLVQLLGTGTDSEPCGLSDLRSIMTKPGCSGYCLSHQLLFFLWARMRGCTQGPLQQSQD
YINLFCANMMDLNRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILSWQKQQEGCFG
EPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFQFSCLILP
SSWDYRSVPYLANFYIFLVETGFHHVAHAGLELLISRDPPSTSQSQVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCATGGCGGGCGCTCTGTGGGGCTTCTTCCCGTCCTGCTGCTG
CTGCTGCTATCBBBBBGTCCAGAGCTGGAGGTGCCGGGGCTGCTGCTGAGGGATCGGAGG
GAGTGGGGTCCGCATAGGAGATCGCTCAAGATTGAGGGCGTGCAGTTGTTCCAGGGTGAAGC
CTCAGGACTGGATCTCGCGGCCGAGTGCTGGTAGACGGAGAAGAGCACGTCGGTTCCCTTAAG
ACAGATGGGAGTTTGTGGTTCATGATATACTTCTGGATCTTATGTAGTGGAAAGTTGTATCTCC
AGCTTACAGATTGATCCCCTCGAGTGGATATCACTTCGAAAGGAAAATGAGAGCAAGATATG
TGAATTACATCAAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCAAATGAAATCTTCAGGT
CCACCTCTTACTTTATTAAAAGGAATCGTGGGCTGGACAGACTTCTAATGAACCCAATGGT
TATGATGATGGTCTCCTTATTGATATTGTGCTCTGCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGAAATGGAGCAGTCATGAATATGCTGAATTCCAACCAGTGGCTGAT
GTTTCTGAGTTCATGACAAGACTCTCTTCAAATCATCTGCAAATCTAGCAGCGGCAGCAG
TAAAACAGGCAAAAGTGGGCTGGAAAAGGAGGTAGTCAGGCCGTCAGAGCTGGCATTGCAC
AAACACGGCAACACTGGTGGCATCCAAGTCTGGAAAACCGTGTGAAGCAACTACTATAAAACTT
GAGTCATCCCGACGTTGATCTTACAACGTGTATGTT
AACTTTTAGCACATGTTTGTACTTGGTACACGAGAAAACCCAGCTTCATTTGTCTGTAT
GAGGTCAATATTGATGTCAGTGAATTAAATTACAGTGTCTTATAGAAAATGCCATTAATAAATTAT
ATGAACTACTATACTTATGTATTTAAATTAAACATCTTAATCCAGAAAATCAAAAAAAA
AAAAAAAAAAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLSDVQSSEVPGAAEGSGGSVGIGDRFKIEGRAVPGVKPQDWISAA
RVLVDGEEHVGFLKTGDSFVVDIPSGSYVEVVSPAYRFDPVRVDITSKGKMRARYVNYIKTSE
VVRLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPVMMMVLPLLI FVLLPKVVNTSDPDMRREME
QSMNMLNSNHELPDVSEFMTRLFSSKSSGKSSGSSKTGKSGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTTCTGTGAGATTGATTTCTCCCCAGTTCCCTGTGGGTCTGAGGGGA
CCAGAAGGGTGAGCTACGTTGGCTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCCAAACAA
GTTTGACATTCCCTGAAATGTCATTCTATCTATTCACTGCAAGTGCCTGCTGTTCCAGGC
CTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGACAGAATAAAGGAGCCACGACCTGTGC
CACCAACTCGCACTCAGACTCTGAACCTCAGACCTGAAATCTCTTCACGGGAGGCTTGGCAGT
TTTCTACTCCTGTGGTCTCCAGATTCAGGCCTAAGATGAAAGCCTTAGTCCTGCCTTCAGC
CTTCTCTGCTGCCACAAACCTCAGGAAATACGAAATGGATTTCTGAGATAACGGGCAGTG
AAGCTGTGTGATGCCACAAACCTCAGGAAATACGAAATGGATTTCTGAGATAACGGGCAGTG
TGCAAGCCAAGATGAAACATTGACATCAGAATCTAAGGAGGACTGAGTCAGACAGACA
AAGCCTGCGAATCGATGCTGCCCTGCGCCATTGCTAAGACTCTATCTGGACAGGGTATTAA
AAACTACCAGACCCCTGACCATTACTCTCCGAAGATCAGCAGCCTGCCAATTCTTCTTA
CCATCAAGAAGGACCTCCGGCTCTCATGCCACATGACATGCCATTGTGGGAGGAAGCAATG
AAGAAATACAGCCAGATTCTGAGTCACTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
TTTGGGGAACTAGACATTCTCTGAATGGATGGAGGAGACAGAATAGGAGGAAAGTGTGCTG
CTGCTAAGAATATTGAGGTCAAGAGCTCCAGTCTCAATACCTGCAGAGGAGGCATGACCCAA
ACCACCATCTTTACTGTACTAGTCTTGCTGGTCACAGTGTATCTTATTTATGCATTACTTG
CTTCCTGCTGATGATTGTCTTATGCATCCCCATCTAATTGAGACCATCTGTATAAGATTT
TGTAATATCTTCTGCTATTGGATATTTATTAGTTAATATTTATTATTTTGCTATTAA
ATGTATTATTTTTACTTGACATGAAACCTTAAAAAAATTCAAGATTATTTATAACCTG
ACTAGAGCAGGTGATGTATTTTATACAGTAAAAAAAAACCTTGAAATTCTAGAAGAGTGG
CTAGGGGGTTATTCATTGTAATTCAACTAAGGACATATTTACTCATGCTGATGCTGTGAGAT
ATTTGAAATTGAACCAATGACTACTTAGGATGGGTTGGAATAAGTTTGATGTGGAATTGCAC
ATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTATCTCCAG
CCAGGAATCCTACACGGCCAGCATGTATTCTACAAATAAGTTCTTGCTACCAAAAAAA
AAAAAAAAAA

FIGURE 138

MRQFPKTSFDISPMSFSIYSLQVPAVPGLTCWALTAEPGWGQNKGATTCATNSHSDSELRPEIF
SSREAWQFFLLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
FSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLRHLLRLYLDdrvFKNYQTPDHYTLRKIS
SLANSFLTIKKDLRLSHAHMTCHGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGGAAGCGCGGCTGCAGAGGGCGAGGCTCAGGTGGGGTGGGTCGGTACCGCATCCAGCC
TAGCGTGTCCACG**T**GGCTGGCTCGGACTTTCGCTACCTGTGCGTAGCGATCGAGGTGCTGG
TAGGGATCGCGGTCTCCCTCGGGATTCTCCCGGCTCCCGTCTGGAGGCTGAGGAGGAA
CACGGAGCGGAGCCCCAGGCCAACCTCGGCTGGAGGCACTCTAAGTGACCACGCTGCC
ACCACCTCTTCAGTAAAGTTGTTATTGTTCTGATAGATGCCCTGAGAGATGATTGTTGTTG
GGTCAAAGGGTGTGAAATTATGCCCTACACAACCTTACCTGTGGAAAAAGGAGCATCTCACAGT
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCCTGAATCAAGGCATTGATGACGGGGAG
CTTCCTGGCTTGTGACGTCACTAGGAACCTCAATTCTCCTGCACGTGCTGGAGACAGTGTGA
TAAGACAAGCAAAGCAGCTGGAAAAAGAATAGTCTTATGGAGATGAAACCTGGGTTAAATTA
TTCCCAAAGCATTTGTGGAATATGATGGAACAAACCTCATTTCGTTGTCAGATTACACAGAGGT
GGATAATAATGTCAGGAGCATTGGATAAAGTATTAAAAGAGGAGATTGGGACATATTAAATCC
TCCACTACCTGGGGCTGGACCACATTGGCACATTTCAGGGCCAACAGCCCCCTGATTGGGAG
AAGCTGAGCAGATGGACAGCGTGTGATGAAGATCCACACCTCACTGCAGTCGAAGGAGAGAGA
GACGCCCTTACCCATTGCTGGTTCTTGTTGACCATGGCATGTCGAAACAGGAAGTCACG
GGGCCTCTCCACCAGGGAGGTGAATACACCTCTGATTAACTCAGTTCTGCTGGTGAAGGAAA
CCCAGGTGATATCCGACATCCAAAGCACGTCCAA**T**AGACGGATGTGGCTGCGACACTGGCGATAGC
ACTTGGCTTACCGATCCAAAAGACAGTGTAGGGAGCTCTATTCCAGTTGTTGGAAGGAAGAC
CAATGAGAGAGCAGTTGAGATTTCACATTGAAATACAGTCAGCTAGTAAACTGTTGCAAGAG
AATGTGCGCTCATATGAAAAAGATCCTGGGTTGAGCAGTTAAAATGTCAGAAAGATTGCA
GAACCTGGATCAGACTGTACTTGGAGGAAAGCATTCAAGGTCTATTCAACCTGGCTCCAAGG
TTCTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTTGTCCCTGAGTCACAGTGGCCAG
TTCTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTTGTCCCTGAGTCACAGTGGCCAG
CTGTCATCTCCGGGTTCTGCTCAGCGTCCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAAGTCCCA
CATTGTTGTCACCTCAGCTGAAAGATTGTCGACTTCTGTCGAGCTGGCCCTCGTGGCTGGCCAGGCT
GCCTTCTGTTACAGACTCTGGTTGAACACCTGGTGTGCCAACGTCGGCAGTGGCCCTGGAC
AGGGGGCCTCAGGGAGGAGCAGTGGAGCAGCCTTATCCCAGGCCCTGGGTGTCGGACACAGGTG
TTCACATCTGTCGTCAGGTCAGATGCCTAGTTCTGGAAAGCTAGGTTCTGCACTGT
CAAGGTGATTGTAAGAGCTGGGGTACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGAA
TCGGACAGCCTCCCAGCAGAGGTGTGGAGCTGCAGCTGAGGGAAAGAAGAGACATCGGCCCTGG
CACTCAGGAGGGTCAAAGAGACTTGGTCGACCAACTCATCCTGCCACCCCCAGAATGCATCT
GCCTCATCAGGTCCAGATTCTTCAAGGGCGACGTTTCTGTTGGAATTCTTAGTCCTGGCC
TCGGACACCTTCATCGTAGCTGGGAGTGGTGGTGGAGGCACTGAGTCAGTGGCAGCAGGACTGTTGGGCCCC
ACCCCAACCTGCACAGCCCTCATCCCCCTCTGGCTGAGCGTCAGAGGCCCTGTCAGACT
CTGACCGAGACACTCACAGCTTGTCACTAGGGCACAGGCTTCTCGGAGGCCAGGATGATCTGT
CCACGCTGACACTCGGGCCATCTGGGCTATGCTCTCTCTGCTATTGAATTAGTACCTAG
CTGCACACAGTATGTTACAAAAAGAATAAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEHGAEPPEPSAGASSNWTTLPPPLF
SKVVIVLIDLALRDDFVFGSKGVKFMPYTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDSVIRQAKAAGKRIVFYGDETWVKLFPKHFVEYDGTTSFVSDYTEVDNNV
TRHLDKVLKRGDWDLILHYLGLDHIGHISGPNSPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGAAGCCTCCAGTTATCGTGACGCACCTGAAAGTCAGAGACTACTGCCCTACA
GAAAGTTACTAGTGCCCTAAAGCTGGCGTGGCACTGATGTTACTGCTGCTGTTGGAGTACA
ACT
TCCCTATAGAAAACA
ACTGCCAGCACCTTAAGACCAC
TCACACCTTCAGAGTGAAGAACTTAAAC
CCGAAGAAATT
CAGCATT
CATGACCAGGAT
CACAAAGTACTGGTCTGGACTCTGGGAATCTCAT
AGCAGTTCCAGATAAAA
ACTACATACGCCAGAGATCTCTTGCA
TTAGCCTCATCCTTGAGCT
CAGCCTCTGCGGAGAAAGGAAGTCCGATT

FIGURE 142

MLLLLLEYNFPIENNQHLKTTHTFRVKNLPKKFSIHQDHKVLVLDGNLIAVPDKNYIRPEI
FFALASSLSSASAEGSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMLAAQKESARRPFI
FYRAQVGWNMLESAAHPGWFICTSCNCNEPGVTDKFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 33-36.

N-myristoylation site.

amino acids 50-55, 87-92

Interleukin-1

amino acids 37-182

FIGURE 143

CTAGAGAGTATAGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCCTCCCTTAA
TCCAGGATCCTGTCTTCCTGTCTGTAGGAGTGCCTGTTGCCAGTGTGGGTGAGACAAGTTG
TCCCACAGGGCTGTCTGAGCAGATAAGATTAAAGGGCTGGGTCTGTGCTCAATTAACTCCTGTGGG
CACGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCATTGCTGGCCTGCCGTGG
AAGGGAGGTCTGTCTGGCGCTGCTGCTGCTCTCTTAGGCTCCCAGATCCTGCTGATCTATGC
CTGGCATTCCACGAGCAAAGGGACTGTGATGAACACAATGTCATGGCTCGTTACCTCCCTGCCA
CAGTGGAGTTGCTGTCCACACATTCAACCAACAGAGCAAGGACTACTATGCCTACAGACTGGGG
CACATCTGAATTCTGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGGAGAACTAGGGTGTGGAAATTGAAAGACGACATTGACAACGTGCCATTCCAAGAAAGCACAG
AGCTGAACAAATACCTCACCTGCTTCTTCACCACAGCACCAGGCCCTGGATGACTCAGTCAGC
CTCCTGAACAAAGACCTGCTTGGAGGGATTCCACTTGAGTGAAACCCACTCACAGGCTTGTCCATGT
GCTGCTCCACATTCCGTGGACATCAGCACTACTCTCCTGAGGACTCTCAGTGGCTGAGCAGCT
TTGGACTTGTGTTATCCTATTTGCATGTGTTGAGATCTCAGATCAGTGTGTTAGAAAATCC
ACACATCTGAGCTAATCATGTAGTAGATCATTAAACATCAGCATTAAAGAAAAAAAAAAAAAA
AA

FIGURE 144

MLGLPWKGGLSWALLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP
WMTQFSLLNKTCLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGGCCAAGGTCTGACGCGATGAGGAAG
CACCTGAGCTGGTGGCTGGCACTGTCTGCATGCTGCTCTCAGCCACCTCTGCGGTCCA
GACGAGGGCATCAAGCACAGAACATCAAGTGAACCGGAAGGCCCTGCCAGCACTGCCAGATCA
CTGAGGCCAGGTGGCTGAGAACCGCCGGAGCCTCATCAAGCAAGGCCGCAAGCTCGACATT
GACTTCGGAGCGAGGGCAACAGGTACTACGGAGGCCAACTACTGGCAGTTCCCCGATGGCATCCA
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTGTACCGGCTGCATCAATGCCA
CCCAGGGCGGAACCAGGGGGAGTTCCAGAAGCCAGACAACAAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTTGCTCCCTCAAGCATTGCGAGTTGGAGAGGGCGCAGGACT
TCGGGTACCATGCAACCAGCCAGTGCTCTGCCTTGATCTGGCTATGGTCAATGGTCAAA
AAGCTTGCAGGAGGCTGGCAGTACAGAGCCAGCAGCGAGCAAATCTGGCAAGTGACCCAGCT
CTTCTCCCCAAACCCACGCGTGTCTGAAGGTGCCAGGAGCGGGCAGTCGCACTCGCACTGCAA
TGCGCTCCCACGTATGCGCCCTGGTATGCGCTCGCTCTGATAGATGGGGACTGTGGCTCT
CCGTCACTCCATTCTCAGCCCCAGCAGCGTCTGGCACACTAGATTAGTAGAAATGCTTGAT
GAGAAGAACACATCAGGCAGTGCAGGCCACCTGCTTCACAGTACTTCCAACACTTGTAGAGGTAG
GTGTATTCCGTTTACAGATAAGGAAACTGAGGCCAGAGAGCTGAAGTACTGCACCCAGCATC
ACCAGCTAGAAAGTGGCAGGCCAGGATTCAACCCCTGGCTGTCTAACCCAGGTTCTGCTCT
GTCCAATTCCAGAGCTGTCTGGTGTACTTATGTCTCACAGGGACCCACATCCAAACATGTAT
CTCTAAATGAAATTGTGAAAGCTCCATGTTAGAAATAATGAAAACACCTGA

FIGURE 146

MRKHLSWWWLATVCMLLFSHLSAVQTRGIKHRIKWNRKALPSTAQITEAQVAENRPGAFIKQGRK
LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNKLHQO
VLWRLVQELCSLKHCFLERGAGLRTMHQPVLLCLLALIWLVMVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCAAAGGGCTGGGATTATAGGCGTGACCACCATGTCCTGGTCCAGAGTCTCATT
CCTGATGATTTATAAGACTCAAAGAAAACTCATGTTCAGAAGCTCTCTCTCTGGCCTCCCT
CTGTCTTCTTCCCTTTCTTCTTAAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
AAATCTTCATTTGCTTGTCAAGTGGGTAGGTCACTGAGTCCTAGTTTATTGGAAATT
CAACTTCAGATTCAAGGGGTACATGTGAAGGTTGTTATGAGTATATTGCATGCTGAGG
TTTGGGGT

FIGURE 148

MFRSSLLFWPPPLCLLSLFLLLILISSIYSESCKLEIFHFACQWGRSLSLSFYFLKFQLSDGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

© 2002 Elsevier Science Ltd. All rights reserved.
0045-2359/\$ - see front matter

FIGURE 149

GTCTCCCGCTCACAGGAACCTCAGCACCCACAGGGCGGACAGCGCTCCCTCTACCTGGAGACTTGAC
TCCCGCGCCCCAACCTGCTTATCCCTGACCCTGAGTGTCAAGAGATCCTGCAGCCGCCAGTCC
CGGCCCCCTCTCCGCCAACCCCACCCACCCTCCTGGCTTCTCTGTTTTACTCCTCCTTCATTCTATA
ACAAAAGCTACAGCTCAGGAGCCAGCGCCGGCTGTGACCCAAGCCAGCGTGGAGAAATGGGGTT
CCTCGGGACGGCACTGGATTCTGGTGTAGTGTCCCATTCAAGCTTCCCAAACCTGGAGGAA
GCCAAGACAAATCTCTACATAATAGAGAATTAGTGAGAAAGACCTTGAATGAACAGATTGCTGAA
GCAGAAGAAGACAAGATTAAAAACATATCCTCCAGAAAACAAGCCAGGTAGAGCAACTATTCTT
TGTGATAACTTGAACCTGCTAAAGGCAATAACAGAAAAGGAAAAATTGAGAAAGAAAGACAATCTA
TAAGAACGCTCCCCACTTGATAATAAGTGAAATGTGGAGATGTTGATTCAACCAAGAATCGAAAATG
ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAAATTCAAGATGATCCAGATGGTCTTC
TCAACTAGACGGACTCCTTAACCGCTGAAGACATTGTCATAAAATCGCTGCCAGGATTATGAG
AAAATGACAGAGGCCGTGTTGACAAGATTGTTCTAAACTACTTAATCTGGCCTTATCACAGAAAGC
CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTACAAAATTAATCTCAAAGGAAGCCAACAA
TTATGAGGAGGATCCAATAAGCCCACAAGCTGGACTGAGAATCAGGCTGGAAAATACCAGAGAAAG
TGACTCCAATGGCAGCAATTCAAGATGGCTTGCTAAGGGAGAAACGATGAAACAGTATCTAACACA
TTAACCTTGACAAATGGCTGGAAAGGAGAACTAAAACCTACAGTGAAAGACAACCTTGAGGAACTCCA
ATATTTCCAAATTCTATGCGCTACTGAAAAGTATTGATTGAGAAAAGAAGCAGAGGAAAGAAA
CACTGATTACTATCATGAAAACACTGATTGACTTGTGAAGATGATGGTGAATGCTTCAAGACCAAAAGCT
CCAGAAGAAGGTGTTCCCTACCTTGAAAATTGGATGAAATGATTGCTTCAAGACCAAAAGCT
AGAAAAAAATGCTACTGACAATATAAGCAAGCTTTCCAGCACCCTCAGAGAAGAGTCATGAAGAAA
CAGACAGTACCAAGGAAGAAGCAGCTAACAGATGGAAAAGGAAATATGGAAGCTTGAAGGATTCCACAAAA
GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAACAGAAGCCTATTGGAAGC
CATCAGAAAAAAATATTGAATGGTTGAAGAAACATGACAAAAAGGGAAATAAGAAGATTATGACCTTT
CAAAGATGAGAGACTTCATCAATAAACAGCTGATGTTATGTGGAGAAAGGCATCCTTGACAAGGAA
GAAGCCGAGGCCATCAAGCGCATTATGAGCAGCTGTAAAATGGCAGGAAAGATCCAGGAGTCTTCAA
CTGTTTGAGAAAACATAATATAGCTTAAACACTCTAATTCTGTGATTAAAATTGGTACCCAGGAA
GTTATTAGAAGTGTGAATTACAGTAGTTAACCTTTACAAGTGGTAAAACATAGCTTCTCC
CTAAAAACTATCTGAAAGTAAAGTGTATGAAAGCTGAAAAAAAAAAAAAAAAAA

FIGURE 150

MGFLGTGTWILVLVLPQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG
QSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPLTAE DIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDPNKPTSWTENQAGK IPEKVTPMAAIQDGLAKGENDET VSNTLTLTNGLE
RRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISPEEGV
SYLENLDEMIALQTKNKLEKNATDNISKLFPAPSEKSHEETDSTKEAAKMEKEYGSLKDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKGKGNKEDYDLSKMRDFINKQADAYVEKGILDK
EEAEAIIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCATACACCCTGTGGAGCTCAA
GATGGTCCCTGAGTGGGGCCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTTATCTGC
ATAATAACCAGCTTCTAGCTGGAGGGCTGCATGCAGGAAAGGTCAATTAAAGGTGAAGAGAGTCAGC
GTGGTCCCCAATCGGTGGCTGGATGCCAGCCTGTCCCCGTCATCTGGGTGTCAGGGTGGAAAG
CCAGTGCCGTATGTGGGTGGGCAGGGCCACTCTAACACTAGAGCCAGTGAACATCATGG
AGCTCTATCTGGTGCAGGAAATCCAAGAGCTTCACCTCTACCGGGGACATGGGGCTCACC
TCCAGCTCGAGTCGGCTGCCTACCCGGCTGGTCTGTGCACGGTGCCTGAAGCCGATCAGCC
TGTCAAGACTCACCCAGCTCCCGAGAATGGTGGCTGGAAATGCCCATCACAGACTCTACTTCC
AGCAGTGTGACTAGGGCAACGTGCCCCCAGAACTCCCTGGGCAGAGCCAGCTCGGGTGGGGT
GAGTGGAGGAGACCCATGGCGGACAATCACTCTCTGCTCTCAGGACCCCACGTCTGACTTAG
TGGGCACCTGACCACTTGTCTCTGGTCTCCAGTTGGATAAAATCTGAGATTGGAGCTCAGT
CCACGGTCCTCCCCACTGGATGGTCTACTGCTGTGGAAACCTTGTAAAACCAGTGGGGTAAA
CTGGGAATAACATGAAAGATTCTGTGGGGTGGGGTGGGGAGTGGTGGGAATCATTCTGCT
TAATGGTAACGTGACAAGTGTACCCCTGAGCCCCGCAGGCCAACCCATCCCCAGTTGAGCCTTATA
GGGTCACTAGCTCCACATGAAGTCCTGTCACTCACCCTGTGCAGGAGAGGGAGGTGGTCATA
GAGTCAGGGATCTATGGCCCTTGGCCCAGCCCCACCCCTTCAATCTGCCACTGTCA
TGCTACCTTCTATCTCTCCCTCATCTTGTGGCATGAGGAGGTGGTGTAGAA
GAAATGGCTCGAGCTCAGAAGATAAAAGATAAGTAGGGTATGCTGATCCTCTTTAAAAACCAA
GATACAATCAAATCCCAGATGCTGGTCTCTATTCCCATGAAAAAGTGCCTCATGACATATTGAGA
AGACCTACTTACAAGTGGCATATATTGCAATTATTAAATTAAAAGATAACCTATTATATT
TCTTTATAGAAAAAGTCTGGAGAGTTACTCAATTGTAGCAATGTCAGGGTGGTGGCAGTAT
AGGTGATTCTTTCTTTAATTCTGTTAATTATCTGTATTCTCAATTCTTCTAATGAGATGA
ATTCCCTGTATAAAAATAAGAAAAGAAAATTATCTGAGGTAAAGCAGAGCAGACATCATCTGA
TTGTCCTCAGCCTCCTCCAGAGTAAATTCAAAATGAAATCGAGCTCTGCTCTGGTGG
TTGTAGTAGTGTAGCAAGGAAACAGATCTCAGCAAAGCCACTGAGGAGGGCTGTGAGTTGT
GTGGCTGGAAATCTGGGTAAAGGAACCTAAAGAACAAAATCATCTGTAATTCTTCTAGAAG
GATCACAGCCCCCTGGGATTCAAGGCATTGGATCCAGTCTCTAAGAAGGCTGCTGTACTGGTGA
ATTGTGCCCCCTCAAATTCAACATCCTCTGGAAATCTCAGTCTGTGAGTTATTGGAGATAAG
GTCTCTGAGATGTAGTTAGTAAGACAAGTCAGTGGATGAAGGTAGACCTAAATTCAATAT
GACTGGTTCTGTATGAAAGGAGAGGACACAGAGACAGAGGAGACGCCGGGAAAGACTATGTA
AAGATGAAGGCAGAGATCGGAGTTTGCAAGCACAAGCTAAGAAACACCAAGGATTGTCACCA
ATCAGAAAGCTGGAGAGGCAAGAAGAAATCTTCCCTAGAGGCTTGTAGGGATAACGGCTCTG
CTGAAACCTTAATCTCAGACTTCCAGCCTCTGAACGAAGAAGAAATAATTCCGGCTTTAA
GCCACCAAGGATAATTGGTACAGCAGCTCTAGGAAACTAATACAGCTGCTAAATGATCCCTG
CTCCTCGTGTTCACATTCTGTGTGTCCTCCACATGACCAAGTGTGTTGTGACCAA
TAGAATATGGCAGAAGTGTGATGGCATGCCACTTCAAGATAGGTATAAAAGACACTGCAGCTC
TACTTGAGCCCTCTCTGTCACCCACCGCCCCAATCTACATCTGGCTCACTCGCTCTGGGG
AAGCTAGCTGCCATGCTATGAGCAGGGCTATAAGAGACTTAGTGGTAAAAATGAAGTCTCCT
GCCACAGCCACATAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATCGATTTGTTGTTT
AAGTTGCTCAGTTGGCTAACTGTGAGATAATATGCAGCAATAGATAATAATGCAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEEISVVPNRWLDAASLSPVILGVQGGS
QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP
VRLTQLPENGGWNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTCCCCAGTCACCAAGTTGCTCGAGTTAGAAATTGCTGCAATGGCCGC
CCTGCAGAAATCTGTGAGCTTCCCTATGGGACCCGGCCACCAGCTGCCTCCCTCTGG
CCCTCTGGTACAGGGAGGAGCAGCTGCCCATCAGCTCCACTGCAGGCTTGACAAGTCCAAC
TTCCAGCAGCCCTATATCACCAACCGCACCTCATGCTGGCTAAGGAGGCTAGCTGGCTGATAA
CAACACAGACGTTCGTCTATTGGGAGAAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACCTCACCTGAAGAAGTGTGCTGTTCCCTCAATCTGATAGGTT
CAGCCTTATATGCAGGGAGGTGGTGCCTCTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAGCTGAAGGACACAGTGAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTGCTGTTATGCTCTGAGAAAT
GCCTGCATTTGACAGAGCAAAGCTGAAAATGAATAACTAACCCCCCTTCCCTGCTAGAAATAA
CAATTAGATGCCCAAAGCGATTTTTAACAAAAGGAAGATGGGAAGCCAACCTCCATCATG
ATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAACCAATGCCACTTTGTTATA
AGACCAGAAGGTAGACTTCTAACGATAGATATTATTGATAACATTCACTGTAACGGTGTTC
TATACACAGAAAACAATTATTAAATAATTGTCTTTCCATAAAAAAGATTACTTCCAT
TCCTTAGGGAAAAACCCCTAAATAGCTCATGTTCCATAATCAGTACTTTATTTATAAAA
TGTATTATTATTATAAGACTGCATTTATTTATCATTATTAATATGGATTATTAT
AGAAACATCATTGATATTGCTACTTGAGTGTAAAGGCTAATATTGATATTATGACAATAATTAT
AGAGCTATAACATGTTATTGACCTCAATAAACACTGGATATCCC

FIGURE 154

MAALQKSVSSFLMGTIATSCLLLALLVQGGAAAPISSHCRLDKSNFQQPYITNRTFMLAKEASL
ADNNNTDVRILIGEKLFHGVSMSERCYLMQVLNFTLEEVLFHQSDRFQPYMQEVVPFLARLSNRSL
TCHIEGDDLHIQRNVQKLKDVTKKLGESGEIKAIGELLLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCGT
CAGTCAGTCCCCGACTTGTGACTGAGTGTGCAGTGCCCAGCATGTACCAGGTCA GTGCAGAGGG
TGCCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGC
TGCCAGGTTGGGCTGGGGCCAAGTGGAGT GAGAAACTGGGATCCCAGGGGAGGGTGCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCATTAGCCTTTCTACAGGTGGTTGCAT
TCTTGGCAATGGTCA TGGGAACCCACACCTACAGCCACTGGCCAGCTGCTGCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCCTCCCTAGAGCCTGCTAG
GCCCAACCGCCACCCAGAGTCCTGTAGGGCCAGTGAAGATGGACCCCTAACAGCAGGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCCTCCCCAGGACCTGTACCACGCCGT
TGCCTGTGCCGCACTGCGTCAGCCTACAGACAGGCTCCACATGGACCCCCGGGCAACTCGGA
GCTGCTTACCAACCAACAGACTGTCTTCTACAGGCGGCCATGCCATGGCGAGAAGGGCACCCACA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTTAGCTTGTGTGTGCGGGCCCGT
GTGATGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTTTGGAAACCTGGAGCCAGGTGTACA
ACCACTGCCATGAAGGGCCAGGATGCCAGATGCTTGGCCCTGTGAAGTGTCTGGAGCAG
CAGGATCCCAGGACAGGATGGGGCTTGGGAAAACCTGCACCTCTGCACATTGAAAAGAG
CAGCTGCTGCTTAGGCCGCCAGCTGGTGTCCCTGTCA TTCTCAGGAAAGGTTTCAA
GTTCTGCCATTCTGGAGGCCACCACTCCTGTCTTCCCTTTCCATCCCCGTACCCCTG
GCCAGCACAGGCACTTCTAGATATTCCCCCTGCTGGAGAAGAAAGAGCCCTGGTTTATT
TGTTGTTACTCATCACTCAGTGAGCATCTACTTGGGTGCATTCTAGTGTAGTTACTAGTCTT
TTGACATGGATGATTCTGAGGAGGAAGCTGTTATTGAATGTATAGAGATTATCAAATAATAT
CTTTATTAATGAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGHTYSHWPSCCP SKGQDTSEELL RWSTVPVPPLEPA
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTHKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCGGATGTGGCTCGTGCTGCCTAACGCCGCGCGTGTGCAGGAGCGCCGTACCCCCGAGAGCC
GACCGTTCAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACACTAGTGTGCAACAGGGGACTATTCA
ATTTTGATGAATGTAAGCTGGGACTCCGGCAGATGCCAGCATCCGCTTGTGAAGGCCACCAA
GATTTGTGTGACGGCAAAAGCAACTTCCAGTCCTACAGCTGTGTGAGGTGCAATTACACAGAGG
CCTTCCAGACTCAGACCAGACCCCTGGTGTAAATGGACATTTCTACATCGGCTTCCCTGTA
GAGCTGAACACAGTCTATTTCATTGGGCCATAATTCCTAAATGCAAATATGAATGAAGATGG
CCCTTCCATGTCTGTGAATTTCACCTCACAGGCTGCCTAGACCACATAATGAAATATAAAAAAA
AGTGTGTCAAGGCCGAAGCCTGTGGGATCCGAACATCACTGCTTGTAAAGAAGAATGAGGAGACA
GTAGAAGTGAACCTCACACCCTCCCCTGGAAACAGATACTGGCTTATCCAACACAGCAC
TATCATGGGTTTCTCAGGTGTTGAGCCACACCAGAAAGAACAAACGCGAGCTTCAGTGGTGA
TTCCAGTGACTIONGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTTCTACTTGTGGC
AGCGACTGCATCCGACATAAAGAACAGTTGTGCTCTGCCACAAACAGGCGTCCCTTCCCTCT
GGATAACAACAAAAGCAAGCCGGAGGCTGGCTGCCTCTCCTCTGCTGTCTGCTGGTGGCCA
CATGGGTGCTGGTGCAGGGATCTATCTAAAGTTCTGGTTACCCATCTGAAATATGTTCCA
TCTACCACCAACTACTGCCCTTCAATTAAAGTTCTGGTTACCCATCTGAAATATGTTCCA
TCACACAATTGTTACTTCACTGAATTCTCAAAACCATTGCAAGAAGTGAAGTCATCCTGAAA
AGTGGCAGAAAAGAAAATAGCAGAGATGGGTCCAGTGCAGTGGCTGCCACTCAAAGAAGGC
GCAGACAAAGTCGCTTCCCTTCCATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAA
GAGCGAGGGCAGTCCAGTGAGAACTCTCAAGACCTTCTCCCCCTGCCTTAAACCTTCTGCA
GTGATCTAAGAAGCCAGATTCTGACAAATACGTGGTGGCTACTTAGAGAGATTGATACA
AAAGACGATTACAATGCTCTCAGTGTCTGCCCAAGTACCACTCATGAAGGATGCCACTGCTTT
CTGTGCAAGAACCTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAGATCACAAGCCTGCCACG
ATGGCTGCTGCCCTGTAG

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLIPGDLRDLRVEPVTTSVATGDYSILMNWSW
LRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIP
NANMNEGDGPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKNEETVEVNFTTPLGNRYMALIQH
STIIGFSQVFEPHQKKQTRASVVIPVTGDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPQTGVPFPLDNNK
SKPGGWLPLLLSSLIVATWVLVAGIYLMWRHERIKTSFSTTLLPPIKVLVVYPSEICFHHTICYFTEFL
QNHCRSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLLSNDVNSVCDGTCGKSEGSPSENSQDLFPLA
FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVCPKYHLMKDATAFCAELLHVVKQQVSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
- 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

FIGURE 159

AGCCACCAGCGCAACATGACAGTGAAGACCCTGCATGGCCCAGCCATGGTCAGTACTTGCTGCT
GTCGATATTGGGGCTTGCCTTCTGAGTGAGGCAGCTCGGAAATCCCAAAGTAGGACATA
CTTTTTCCAAAAGCTGAGAGTTGCCGCCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
ATCATCAATGAAAACCAGCGCTTCCATGTCACGTAACATCGAGAGCCGCTCCACCTCCCCCTG
GAATTACACTGTCATTGGGACCCCAACCAGTACCCCTCGGAAGTTGTACAGGCCAGTGTAGGA
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTTCCATCCAGCAA
GAGACCCCTGGTCGTCCGGAGGAAGCACCAAGGCTGCTCTGTTCTTCCAGTTGGAGAAGGTGCT
GGTGACTIONTGGCTGCACCTGCGTCACCCCTGTCATCCACCATGTGCAGTAAGAGGTGCATATCC
ACTCAGCTGAAGAAG

FIGURE 16O

MTVKTLHGPAMVKYLLSILGLAFLSEAAARKIPKVGHTFFQKPESC PPPVPGGSMKLDIGIINEN
QRVSMSRNIESRSTSPWNYTVTWDPNRYPSEVVQAQCRLGCINAQGKEDISMNSVPIQQETLVV
RRKHQGCSVFQLEKVLVTVGCTCVTPVIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

ACACTGGGCCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGGAGTCAGGACTCCCAGG
ACAGAGAGTGACAAACTACCCAGCACAGCCCCCTCGCCCCCTCTGGAGGCTGAAGAGGGATT
CAGCCCCCTGCCACCCACAGACACGGGCTGACTGGGTGTCGCCCCCTTGGGGGGGGCAGCAC
AGGGCCTCAGGCCTGGGTGCCACCTGGCACCTAGAAGATGCTGTGCCCTGGTTCTGCTGCT
TGGCACTGGGCCGAAGCCCAGTGGCTCTTCTCTGGAGAGGCTGTGGGCCAGGACGCTACC
CACTGCTCTCGGGCCTCTCTGCCCTCTGGACAGTGACATACTCTGCCCTGCCGGGACAT
CGTGCCTGCTCCGGGGCCCGTGTGGCGCTACGCACCTGCAGACAGACTGGTGTGAGGTGCC
AGAAGGAGACCGACTGTGACCTCTGCTGCGTGTGGCTGTCCACTGGCGTGCATGGCACTGG
GAAGAGCCTGAAGATGAGGAAAAGTTGGAGGAGCAGCTACAGGGTGGAGGAGCAGCT
TGCCTCTCCAGGCCAACAGTGTGCTCTCCAGGCCACCTACTGCGCTCTGC
TGGAGGTGCAAGTGCCCTGCTGCCCTTGTGCAGTTGGTCACTGTGCGCTACT
TGCTTCGAGGCTGCCCTAGGGAGTACGAATCTGGCTACTACGCCAGGTACGAGAA
GGAACATCACCACACAGCAGCTGCCCTGCCCTGCCCTGCCCTGCCCTGCC
ACGTGCATCTGGTTCTGAATGTCTCTGAGGAGCAGCACTTGGCCCTCTCCCTGTACTGGAAATCAG
GTCCAGGGCCCCCCTAAACCCCGTGGCACAAAAACCTGACTGGACCCAGATCATTACCTTGA
CCACACAGACCTGGTCTCCCTGCCCTGTATTCAGGTGTGGCCTGTGAAACCTGACTGGTCTAGGA
CGAACATCTGCCCTTCAGGGAGGACCCCCGCGCACACAGAACCTGTGGCAAGCCGCCACTG
CGACTGCTGACCCCTGAGAGCTGGCTGCTGGACGCACCGTGCTGCTGCCCGAGAACGGGACT
GTGCTGGGGCTCCGGGTGGGACCCCTGCCAGCCACTGGTCCACCCGTTCTGGAGAACG
TCACTGTGGACAAGGTTCTGAGTTCCCATGGCTGAAAGGCCACCTAACCTCTGTGTTCAAGGTG
AACAGCTGGAGAAGCTGCAGCTGCAGGAGTGTGCTTGTGGCTGACTCCCTGGGCCCTCAAAGA
CGATGTGCTACTGTTGGAGACAGCAGGAGGCCCAAGGACAACAGATCCCTGTGCCCTGGAAACCCA
GTGGCTGACTTCACCTACCCAGCAAAGCCTCACAGGGCAGCTGCCCTGGAGAGACTTACTA
CAAGACCTGCAGTCAGGCCAGTGTCTGCAGCTATGGGACGATGACTTGGAGCGCTATGGGCTG
CCCCATGGACAAATACATCCACAGCGTGGCCCTCGTGTGGCTGGCCTGCCTACTTTGCCG
CTGCGTTCCCTACCTCTCTCAAAAGGATCAGGCAAAAGGGTGGCTGAGGCTCTTGA
CAGGACGTCCGCTCGGGGGCGGCCAGGGCAGGCCGCGCTCTGCTCTACTCAGCGATGA
CTCGGGTTTCGAGGCCCTGGTGGGCCCTGGCGTCTGCCCTGTGCCAGCTGCCGTGCGCTGG
CCGTAGACCTGTGGAGGCCCTGCTGAGACTGAGCGCGCAGGGGCCCTGGCTTGGTTCA
CGGGGCCAGACCCCTGAGGCCCTGGTGGCTTCTGCTCTCTCCCGTGGCTGGCG
GTGCAGCGAGTGGCTACAGGATGGGTGTCGGGCCCGGGCGCACGGCCCGCACAGCGCTTCC
GCGCCTCGCTCAGCTGCCTGCTGCCGACTTCTGCAAGGGCCGGCGCCCGAGTCAGTGGGG
GCCCTGCTCAGGGCTGCTCCACCCGGACGCCGTACCCGCCCTTTCGCACCGTGCCTGCT
CACACTGCCCTCCCAACTGCCAGACTTCTGGGGGCCCTGCAGCAGCCTCGGCCCGCGTCCG
GGCGCTTCAAGAGAGAGCGGAGCAAGTGTCCGGGCCCTCAGCCAGCCCTGGATAGCTACTTC
CATCCCCGGGGACTCCCGCGCCGGGACGCCGGGGTGGGACCCAGGGGGGGGACCTGGGGCGGGG
CGGGACTTAAATAAAGGAGACGCTGTTTCTAAAAAA

FIGURE 162

MPVFWFLLSLALGRSPVVLSLERLVGPQDATHCSPGLSCRWLDSILCLPGDIVPAPGPVLAPTHLQTELVLRCQKETDCDLCLRVAVHLAVGHWEEPEDEEKFGGAADSGVVEPRNASLQAQVVLFSQAYPTARCVILLEVQVPAALVQFGQSVGSVVYDCFEAALGSEVRIWSYTQPRYEKELNHTQQLPALPWLNVSADGDNVHVLVNVSEEQHFGLSLYWNQVQGPPKPRWHKNLTGPQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQNLWQAARLRLLLTQSLLDAPCSLPAEAALCWRAPGGDPCQPLVPPLSWENVTVDKLEFPPLLKGHPNLCVQVNSSEKQLQECWLADSLGPLKDDVLLLETTRGPQDNRSLCALEPSGCTSLSKASTRAARLGEYLLQDLQSQCQLQLWDDDLGALWACPMDKYIHKRWALVWLACLLFAAALSLLKKDHAKGWLRLLKQDVRSGAARGRAALLLYSADDSGFERLVLGALASALCQLPLRVAVDLWSRRELSAQGPVAWFHAQRQTLQEGGVVVLLFSPGAVALCSEWLQDGVSAGPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLLHPDAVPALFRTPVVFTLPSQLPDFL GALQQPRAPRSGRILQERAEQVSRALQPALDSYFHPPGTPAPGRGVGPGAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251, 334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617, 692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGCCAGCCCCGATGAGGACGCTGCTGACCACITTGACTGTGGATCCCTGGCT
GCTCACGCCCCCTGAGGACCCCCTCGGATCTGCTCCAGCAGTGAAAATTCCAGTCCAGCAACTTTGA
AAACATCCTGACGTGGGACAGCGGGCCAGAGGGCACCCTAGACACGGTCTACAGCATCGAGTATA
AGACGTACGGAGAGAGGGACTGGGTGGAAAGAAGGGCTGTGCAGGGATCACCCGAAGTCCCTGC
AACCTGACGGTGGAGACGGCAACCTCACGGAGCTACTATGCCAGGGTACCCGCT
GTCAGTGGGGAGGGCCGGTCAGCCACCAAGATGACTGACAGGTTCAGCTCTGCAGCACACTAC
CCTCAAGCCACCTGTGACCTGTATCTCAAAGTGAGATGATTAGATGTTCATCTA
CCCCCACGCCAATCGTGCAGGCATGGCACCCTGAACCTGGAAAGACATCTTCCATGACCTG
TTCTACCAACTAGAGCTCCAGGTCAACCGCACCTACCAAATGCACCTGGAGGGAGCAGAGAGA
ATATGAGTTCTCGGCCGTGACCCCTGACACAGAGTCTTGGCACCACATGATTGCGTTCCA
CCTGGGCAAGGAGAGTCCCCCTACATGTGCCAGTGAAGACACTGCCAGACCGGACATGGACC
TACTCCCTCTCCGGAGCCTCTGTCTCATGGGCTCCTCGTGCAGTACTCTGCTACCTGAG
CTACAGATATGTCACCAAGGCCCTGCACCTCCCAACTCCCTGAACGTCAGCCAGTCCGACTT
TCCAGCCGTGCCCTCATCCAGGAGCACGTCTGATCCCTGTTGACCTCAGCGGGCCCCAGC
AGTCTGGCCAGCCTGTCCAGTACTCCAGATCAGGGTGTCTGGACCCAGGGAGCCCGAGGAGC
TCCACAGCGCATAGCCTGTCCGAGATCACCTACTTAGGGCAGCCAGACATCTCATCCCTCAGC
CCTCCAACTGTGCCACCTCCCCAGATCCTCTCCCCACTGTCTATGCCCAAACGCTGCCCTGAG
GTCGGGCCCCATCTATGCACTCAGGTGACCCCCGAAGCTCAATTCCATTCTACGCCACACA
GGCCATCTAAGGTCCAGCCTCCCTATGCCCTCAAGCCACTCCGGACAGCTGGCCTCCCT
CCTATGGGTATCATGGAAGGTTCTGGCAAAGACTCCCCACTGGGACACTTTCTAGTCCTAAA
CACCTAGGCTAAAGGTCAAGCTCAGAAAGAGCCACAGCTGGAGCTGATGTTAGGTGGCT
TTCTCTGCAGGAGGTGACCTCTTGGCTATGGAGGAATCCAAGAGCAAATCATTGACCCAGC
CCCTGGGATTGACAGACAGAACATCTGACCCAAATGTGCTACAGTGGGAGGAAGGGACA
CCACAGTACCTAAAGGGCAGCTCCCTCTCCTCAGTCCCTCAGTCCAGATGAGGGCACCCATGTC
CCTCCCTTGCACACTCTCCGGTCCATGTTCCCTCGGACCAAGGTCCAAGTCCCTGGGCC
TGCTGGAGTCCCTGTGTCTCCAGGATGAAGCCAAGAGCCCAGCCCTGAGACCTCAGACCTG
GAGCAGCCACAGAACTGGATCTCTTCAAGGGCTGGCCCTGACTGTGCAGTGGAGTCTGAGGG
AGGGGAATGGAAAGGCTTGGTGTCTCCCTGTCCCTACCCAGTGTACATCTTGGCTGTCA
ATCCCATGCCTGCCCATGCCACACACTCTGCATCTGCCCTCAGACGGGTGCCCTGAGAGAAC
AGAGGGAGTGGCATGCCAGGGCCCTGCCATGGGTGCCTCTCACCAGAACAAAGCAGCATGATA
AGGACTGCAGGGGGAGCTCTGGGGAGCAGCTGTGAGACAAGCCGTGCTGTCAGGCTG
CAAGGCAGAAATGACAGTGCAGGAGGAATGCAGGGAAACTCCCAGGTCCAGGCCACCTC
CTAACACCATGGATTCAAAGTGTCAAGGAATTGCCCTCCTGCCCTGGCCAGTTTC
ACAATCTAGCTGCCACAGAGCATGAGGCCCTGCCCTCTGTCAATTGTTCAAAGTGGAGAGA
GCCGGAAAAGAACCAAGGCCCTGGAAAAGAACCAAGCAGAAGGAGGCTGGCAGAACAGAAC
ACTTCTGCAAGGCCAGGGCAGCAGGACGGCAGGACTCTAGGGAGGGGTGTGGCTGAGCTCA
TTCCAGGCCAGGCAACTGCCCTGACGTTGACGATTCTAGCTTCAATTCTCTGATAGAAC
GAAATGCAGGTCCACCAAGGGAGACACACAAGCCTTTCTGCAAGGCAGGAGTTCAAGAC
ATCTGAGAATGGCTTGTCAAAGGAAGGTGAGGGCTGTGGCCCTGGACGGTACAATAACAC
TGTACTGATGTCAAACACTTGCAGCTGCTGGCTGGGTTCAAGCCCCATCTGGCTCAAATT
CTCACCAACTCACAAGCTGTGACTTCAAACAAATGAAATCAGTGCAGGAGAACCTCGGTT
ATCTGTAATGTGGGATCATAACACCTACCTCATGGAGGTGTGGTGAAGATGAAAGTCATG
TCTTAAAGTGTAAATAGTGCCTGGTACATGGCAGTGCCAATAACGGTAGCTATTAAAAA
AAAAAAA

FIGURE 164

MRTLLTILTGVSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPETPDTVYSIEYKTYGERDW
VAKKGCQRITRKSCNLTVETGNLTELYYARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVT CIS
KVRSIQMIVHPTPTPIRAGDGHLTLTLEDIFHDLFYHLELQVNRTYQMHLGGKQREYEFFGLTPDT
EFLGTIMICVPTWAKESAPYMCRVKTLPDRTWTYSFSGAFLFSMGFLVAVLCYLSYRYVTKPPAP
PNSLNQVRVLTFQPLRFIQEHLVLI PFDLSGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT
YLGQPDISILOPSNVPPPQI LSPLSYAPNAAPEVGPPSYAPQVTPEAQFPFYAPQAISKVQPSY
APQATPDSWPPSYGVCMEGSGKDSPTGTLS SPKHLRPKGQLKEPPAGSCMLGGLSLQEVTSLAM
EESQEAKSLHQPLGICTDRTSDPNVLHS GEEGTPQYLKQQLPLLSSVQIEGHPMSLPLQPPSGPC
SPSDQGPSPWG LIESLVCPKDEAKSPAPETSDLEQPTELDSLFRGLALT VQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAGTCACCCGGGCCGCGGTGCCACAACATGG
CTGGCGCGCCGGGCTGCTCTCTGGCTGTCGTGGGGCGCTCTGGTGGTCCCGGCCAG
TCGGATCTCAGCCACGGACGGCGTTCTCGGACCTCAAAGTGTGCGGGGACGAAGAGTGCAGCAT
GTTAATGTACCGTGGAAAGCTCTGAAGACTTCACGGGCCCTGATTGTCGTTTGAATTTA
AAAAAGGTGACGTATATGTCTACTACAAACTGGCAGGGGATCCCTGAACTTGGCTGGA
AGTGTGAACACAGTTGGATTTTCAAAGATTGATCAAGTACTTCATAAATACACGGA
AGAAGAGCTACATATTCCAGCAGATGAGACAGACTTTGTCTGCTTGAAGGAGAGATGATT
TTAATAGTTATAATGTAAGAAGAGCTTTAGGATCTTGGAACTGGAGGACTCTGTACCTGAAGAG
TCGAAGAAAGCTGAAGAAGTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTCGGGGGCGTGA
ACTTGACCTGTGCCTGAGCCGAGGCATTCAGAGCTGATTCAGAGGATGGAGAAGGTGCTTCT
CAGAGGCACCGAGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGGCCACCCTCACACCAGCGGT
CCTGCGGCTAACCCCTCAGGGAGTGCAGTCTCGTGGACACTTTGAAGAAATCTGCACGATAA
ATTGAAAGTGCGGGAAGCGAAAGCAGAACTGGCAATAGTTCTCCTGCCTCGGTGGAGCGGGAGA
AGACAGATGCTACAAAGTCTGAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGCTTATT
CATTACAGCAAAGGATTCTGTGGCATCAAATCTAAGTTGTTACAAAGATTGTTTAGTA
CTAAGCTGCCTTGGCAGTTGCATTTTGAGCCAACAAAAATATATTATTCCCTCTAAGTA
AAAAAAAAAAAAAA

FIGURE 166

MAAPGLLFWLFVLGALWWVPQSDLSHGRRFSDLKVCGDEECMSMLMYRGKAEDFTGPDCRFVN
FKKGDDVYVYYKLAGGSLEIWAGSVEHSFGYFPKDLIKVLHKYTEEELHIPADETDFVCFEGGRD
DFNSYNVEELLGSLELEDSPPEESKKAAEVSQHREKSPEESRGRELDPVPEAFRADSEdgeGA
FSESTEGLQGQPSAQESHPTSGPAANAQGVQSSLDTFEEILHDKLKVPGSESRTGNSSPASVER
EKTDAYKVLKTEMSQRGSGQCVIHYSKGFRWHQNLSLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCAGTTGTCAAGAGGCCGGGAAGAGAAGCAAAGCGC
AACGGTGTGGTCCAAGCCGGGCTCTGCTCGCCTCTAGGACATACACGGGACCCCTAACTTC
AGTCCCCAAACCGCGACCCCTCGAAGTCTGAACTCCAGCCCCGACATCCACCGCGGGCACAGG
CGCGCAGGCGCAGGTCCCGCGAAGGCATGCGCGCAGGGGCTGGGCAGCTGGCTCGGC
GGCAGGGAGTAGGGCCCGGCAGGGAGGCAGGGAGGCTGCATATTCAAGAGTCGCGGGCTGCGCCCTG
GGCAGAGGCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGCGATGAGCCGCGTGG
TCTCGCTGCTGCTGGCGCCGCTGCTCTGCGGCCACGGAGCCTCTGCGCCCGCTGGTCAGC
GGCCAAAAGGTGTGTTTGCTGACTCAAGCATCCCTGCTACAAATGGCTACTTCCATGAAC
GTCCAGCCGAGTGAGCTTCAGGAGGCACGCCCTGGTTGAGAGTGAGGGAGGAGTCCTCCTCA
GCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAAACCTGACAAAACCCGGG
ACAGGGATTTCTGATGGTGAATTCTGGATAGGGCTTGGAGGAATGGAGATGGGAAACATCTGG
TGCCTGCCAGATCTACCACTGGTCTGATGGAAGCAATTCCAGTACCGAAACTGGTACACAG
ATGAACCTCCTCGGAAGTGAAAAGTGTGTTGATGATCACCAACCAACTGCCAATCCTGGC
CTTGGGGGTCCCTACCTTACCACTGGATGACAGGTGTAACATGAAGCACAATTATATTG
CAAGTATGAACCAGAGATTAATCCAACAGCCCCTGTTAGAAAAGCCTTATCTACAAATCAACCAG
GAGACACCCATCAGAATGTGGTTGTTACTGAAGCAGGTATAATTCCAATCTAATTATGTTGTT
ATACCAACAATACCCCTGCTCTACTGATACTGGTTGCTTTGGAACCTGTTGCTCCAGATGCT
GCATAAAAGTAAAGGAAGAACAAAATAGTCAAACACCAGTCTACACTGTGGATTCAAAGAGTA
CCAGAAAAGAAAGTGGCATGGAAAGTATAAACTCATTGACTGGTTCCAGAATTGTAATTCT
GGATCTGTATAAGGAATGGCATCAGAACAAATAGCTGGATGGCTGAAATCACAAAGGATCTGC
AAGATGAACTGTAAGCTCCCCCTGAGGCAAATATTAAAGTAATTATGTTATATGTCTATTATTCA
TTTAAAGAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTGCTAAAGGATGCACCCAA
ACTTCAAACCTCAAGCAAATGAAATGGACAATGCAAGATAAAAGTTGTTATCAACACGTCGGGAGTA
TGTGTGTTAGAACAAATTCTTATTCTTACCTTCATAAGTTGTTATCTAGTCATGTAA
TGTATATTGTTAGAACATTACAGTGTGCAAAGTATTACCTTGCATAAGTGTGTTGATAAAA
ATGAACTGTTCTAATATTATTATGGCATCTCATTTCATAACATGCTCTTGATTAAAG
AAACTTATTACTGTTGTCACTGAATTCACACACACACAAATATAGTACCATAGAAAAGTTG
TTTCTGAAATAATTCACTTCACTGGTCTGCTTTGGTCAATGCTAGGAAATCTCTCAGA
AATAAGAAGCTATTCAATTAAAGTGTGATATAAACCTCTCAAACATTACTAGAGGCAAGGAT
TGTCTAATTCAATTGTGCAAGACATGTGCCTTATAATTATTAGCTTAAATTAAACAGATT
TTGTAATAATTGTAACCTTGTAAATAGGTGCAAAACACTAATGCAATTGCAACAAAAGAAG
TGACATACACAATATAATCATATGCTTCAACGTTGCTTATAATGAGAACAGCAGCTCTG
GGGTTCTGAAATCAATGTGGTCCCTCTTGCCCCACTAAACAAAGATGGTTGTTGGGGTTGG
ATTGACACTGGAGGAGCAGATAGTGCAAAGTTAGTCTAAGGTTCCCTAGCTGTTAGCCTCTG
ACTATATTAGTATACAAAGAGGTCAATGTGGTTGAGACCAGGTGAATAGTCATATCAGTGTGGAG
ACAAGCACAGCACACAGACATTAGGAAGGAAAGGAACATCGAAATCGTGTGAAAATGGTTGG
AACCCATCAGTGATCGCATATTCAATTGATGAGGGTTGCTGAGATAGAAATGGTGGCTCCTT
CTGCTTATCTCTAGTTCTTCAATGCTTACGCCCTGTTCTCAAGAGAAAAGTTGTAAC
CTGGTCTTCAATGTCCCTGTGCTCCTTTAACAAATAAGAGTTCTGTTCTGGGGAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVVSLLLGAALLCGHGFCCRVSQKVCADFKHPCYKMAFHELSSRVSFQEARLACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQTSGACPDLYQWSDGSNSQ
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLILVAFGTCCFQMLHKSKGRTKTSPNQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-145, 212-217